داده < کیفی (قابل اندازه گیری نیست) < اسمی / ترتیبی

کمی (قابل اندازه گیری) < پیوسته: اگر اعداد را اختیار کرد می‌توان مابین آن عدد را انتخاب کنم ۱۵ — ۱۶ (۱۵/۳ ، ۱۵/۴)

گسسته: بین اعداد اتفاقی نمی‌افتد ۱۴-۱۵-۱۶-۱۷

— صفر مطلق ندارد / آمار معنی ندارد مثال

(مطلق) → kg ۰ ۱ ۲

gr ۰ ۱۰۰ ۲۰۰

(فاصله‌ای) → $F = \frac{9}{5}C + 32$ درجه حرارت

C ۰ ۱۵ ۴۵

F ۳۲ ۵۹ ۱۱۳

← انجام صفر مطلق نیست

$\frac{45}{15} \neq \frac{113}{59}$

صفر < قرار دادی / مطلق $\frac{45-15}{45-0} = \frac{113-59}{113-32}$

جمعیت آماری ← زیر مجموعه‌ای از مجموعه‌ها هستند که اعضای آن حداقل یک خصوصیت مشترک باشد

جمعیت را با حروف بزرگ N نشان می‌دهیم و اگر با حروف بزرگ نوشته شود یک مقدار عموماً مجهول است

صفت < مشخصه = در تمام افراد جامعه ثابت است

متغیر = از فردی به فرد دیگر در حال تغییر است — مثال: خصوصیات افراد دانشجوها، بچه‌های دانشگاه آزاد تهران غرب ثابت است

متغیر هستند مثل وزن، سن و ....

متغیر تصادفی ← متغیر تصادفی یک صفت است که تغییرات آن از فردی به فرد دیگر قابل پیش بینی است و تصادفی است

سرشماری = در بعضی جاها معنا ندارد مثلاً طول عمر لامپ‌های شرکت الوند و باید نمونه گیری کنیم و سرشماری معنا ندارد

نمونه گیری < طبقه بندی / خوشه‌ای / سیستماتیک / تصادفی

تصادفی ← یعنی شانس انتخاب یکسان به همه چیز بدهیم یعنی شانس انتخاب همه یکسان باشد.

مثال < با جایگزینی (مستقل هست) $\frac{1}{N} \times \frac{1}{N} \times \frac{1}{N}$

بدون جایگزینی (مستقل نیست) $\frac{1}{N} \times \frac{1}{N-1} \times \frac{1}{N-2}$

مثلاً از ۳۰ نفر ۱ نفر را انتخاب $\frac{1}{30}$ و حالا می‌خواهیم دو نفر را انتخاب کنیم $\frac{1}{30}$ (با جایگزینی)

Scanned with CamScanner

مثال: قرعه‌کشی: قرعه‌کشی ۱ نفر را انتخاب و دوباره داخل سبد انداخته می‌شود یعنی انتخاب نفر اول هست (با جایگذاری)
دوباره حق انتخاب دارد

اگر در قرعه‌کشی ۱ نفر را انتخاب و دوباره داخل سبد نیندازیم و نسبت مجموعه نداریم (بدون جایگذاری)

درس ۱.۱۳

آمار توصیفی ← جدول فراوانی‌ها
← نمایش داده‌ها (اشکال آماری)
← شاخص‌های آماری

〈 توزیع جدول فراوانی برای داده‌های گسسته 〉 ←

X (در جمعیت بزرگ، مقادیر متغیر)	f_i فراوانی
x_1 (در جمعیت کوچک، مقادیر متغیر)	f_1
x_2	f_2
⋮	⋮
x_k	f_k

$$\sum_{i=1}^{k} f_i = n$$

مثال: تعداد اولاد خانواده‌ها در یک جامعه

X تعداد	فراوانی f_i	$f_i \times x_i$
۰	۱	۰
۱	۲	۲
۲	۳	۶
۳	۴	۱۲
۵	۲	۱۰
	۱۲ مجموع	۱۲۸ مجموع

پنج سطر داریم پس $k=5$

۰
۱ ۱
۲ ۲ ۲
۳ ۳ ۳ ۳ ← یعنی ۴ نفر بودن از خانواده ۳ نفری
۵ ۵

$$\sum_{i=1}^{5} f_i = f_1 + f_2 + f_3 + f_4 + f_5$$

$$1+2+3+4+2=12$$

* N = جمعیت (با حروف بزرگ) مقدارش مجهول است
* n = نمونه (با حروف کوچک) مقدارش همیشه معلوم است

〈 داده‌های پیوسته 〉

طبقات	f_i	m_i
$L_1 - U_1$	f_1	
$L_2 - U_2$	f_2	
⋮	⋮	
$L_k - U_k$	f_k	

	F_i	$m_i = \frac{L_i + U_i}{2}$	$m_i \times F_i$
۰-۴	۲	۲	۲×۲
۴-۸	۳	۶	۳×۶
۸-۱۲	۵	۱۰	۵×۱۰
۱۲-۱۶	۴	۱۴	۴×۱۴
۱۶-۲۰	۱	۱۸	۱×۱۸

$$\sum_{i=1} F_i = n$$

طول طبقات $I_i = U_i - L_i$

Scanned with CamScanner

فرمول (k) استرجس

$$k = 1 + 3{,}32 \log n$$

	f_i
۰-۴	۱
۴-۸	۳
۸-۱۲	(۰) صفر داریم
۱۲-۱۶	۷
۱۶-۲۰	۲

۵ سر داریم

و سطر آن حذف می‌شود

۲ و به طبقه بالا میدهیم و ۲ به طبقه پایین [illegible]

	f_i
۰-۴	۱
۴-۱۰	۳
۱۰-۱۶	۷
۱۶-۲۰	۲
	۱۳

یکی از خصوصیات داده پیوسته این است

* برای حذف صفر (۰) بطور مساوی به طبقه بالا و طبقه پایین تقسیم می‌کنیم

مثال مرکب دهی جدول را دسته بندی کنید

دسته بندی

	f_i
NL ۱	۳ مرکب
NL ۲	۵
NL ۳	۲
NL ۴	۱
⋮	
NL ۹۸	۱
NL ۹۹	۰
NL ۱۰۰	۱
	۱۵۷

X	f_i	$m_i = \frac{L_i + U_i}{2}$
۱-۱۰	۱۵	۵٫۵
۱۱-۲۰	۱۷	۱۵٫۵
۲۱-۳۰	۱۱	۲۵٫۵
۳۱-۴۰		
⋮		⋮
۹۱-۱۰۰	۸۰	۹۵٫۵
	۱۵۷	

از خصوصیات

این جدول را جدول پیوسته نمی‌گویند

اگر بخواهیم این جدول را پیوسته کنیم می‌شود:

	f_i	m_i
۰٫۵-۱۰٫۵	۱۵	۵٫۵
۱۰٫۵-۲۰٫۵	۱۷	۱۵٫۵
⋮	⋮	
۹۰٫۵-۱۰۰٫۵	۸	۹۵٫۵
	۱۵۷	

نشان دهنده پیوستگی است

مثال کره

۴۹	(۵۰)	۵۱
۴۹٫۵		۵۰٫۵

Scanned with CamScanner

(داده‌ها را پیوسته کردیم)

تصحیح پیوستگی →

x	f_i
۰	۱
۱	۳
۲	۵
۳	۲
۴	۱
	۱۲

صفر داریم نه ← ۰

	f_i	m_i	$m_i \times f_i$
۰/۵ - ۰/۵	۱	۰	۱×۰ = ۰
۰/۵ - ۱/۵	۳	۱	۳×۱ = ۳
۱/۵ - ۲/۵	۵	۲	۵×۲ = ۱۰
۲/۵ - ۳/۵	۲	۳	۲×۳ = ۶
۳/۵ - ۴/۵	۱	۴	۱×۴ = ۴
	۱۲		

نشان دهنده پیوستگی است

تصحیح پیوستگی اینست ← بین دوتا متغیر چقدر فاصله است و تقسیم بر ۲ می‌کنیم بین او ۰/۵ $= \frac{1}{2}$

جدول اول را به تصحیح پیوستگی تبدیل کنید

$$\sum_{i=1}^{k} f_i = n$$

$$\frac{\sum_{i=1}^{k} f_i}{n} = 1$$

$$\sum_{i=1}^{k} \frac{f_i}{n} = 1$$

$r_i \simeq p_i$

$0 \leq r_i \leq 1$

$f_i = r_i \times n$

x	f_i	$r_i = \frac{f_i}{n}$	$F_i = \sum_{j=1}^{i} f_j$	$R_i = \sum_{j=1}^{i} r_j = \frac{F_i}{n}$
x_1	f_1	r_1	$F_1 = f_1$	$R_1 = r_1$
x_2	f_2	r_2	$F_2 = f_1 + f_2$	$R_2 = r_1 + r_2$
⋮			⋮	
x_k	f_k	$r_k = 1 - \sum_{i=1}^{k-1} r_i$	$F_k = \sum_{i=1}^{k} f_i = n$	
	$\sum_{i=1}^{k} f_i = n$	$\sum_{i=1}^{k} r_i = 1$		

x	f_i	r_i	F_i	R_i
۰	۱	$\frac{۱}{۱۲}$	۱	$\frac{۱}{۱۲}$
۱	۳	$\frac{۳}{۱۲}$	۴	$\frac{۴}{۱۲}$
۲	۵	$\frac{۵}{۱۲}$	۹	$\frac{۹}{۱۲}$
۳	۲	$\frac{۲}{۱۲}$	۱۱	$\frac{۱۱}{۱۲}$
۴	۱	$\frac{۱}{۱۲}$	۱۲	$\frac{۱۲}{۱۲} = 1$
	۱۲	۱		

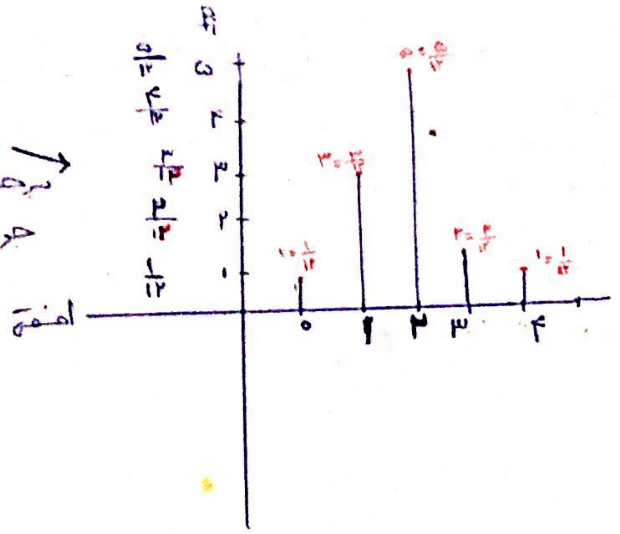
۱ → ۱
۱+۳ = ۴
۱+۳+۵ = ۹
۱+۳+۵+۲ = ۱۱
۱+۳+۵+۲+۱ = ۱۲

نمودار میله‌ای ←

x	f_i	r_i	F_i	R_i
۰	۱	$\frac{1}{12}$	۱	$\frac{1}{12}$
۱	۳	$\frac{3}{12}$	۴	$\frac{4}{12}$
۲	۵	$\frac{5}{12}$	۹	$\frac{9}{12}$
۳	۲	$\frac{2}{12}$	۱۱	$\frac{11}{12}$
۴	۱	$\frac{1}{12}$	۱۲	$\frac{12}{12}=1$
	۱۲	(۱)		

باید ۱ بشود

بر حسب فراوانی جدول مقابل →



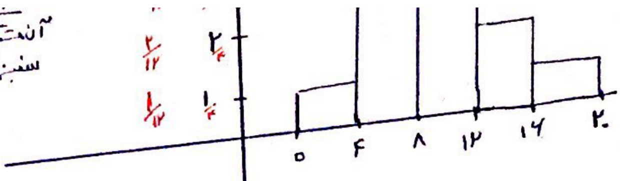
حدود	f_i	r_i	F_i	R_i
۰ - ۴	۱	$\frac{1}{12}$	۱	$\frac{1}{12}$
۴ - ۸	۳	$\frac{3}{12}$	۴	$\frac{4}{12}$
۸ - ۱۲	۵ (مرکز دسته)	$\frac{5}{12}$	۹	$\frac{9}{12}$
۱۲ - ۱۶	۲	$\frac{2}{12}$	۱۱	$\frac{11}{12}$
۱۶ - ۲۰	۱	$\frac{1}{12}$	۱۲	۱۰
	۱۲	۱		

... مستطیل ها خواهد داشت ... مساحت

آن مستطیل دقیقاً برابر است با فراوانی یا فراوانی نسبی آن بازه است

* اگر از فراوانی استفاده کنیم مجموع مساحت مستطیل بدست آمده N

* اگر از فراوانی نسبی استفاده کنیم مجموع مساحت مستطیل بدست آمده یک



⟨ داده پیوسته ⟩

فراوانی

مساحت طبقه i ، دقیق ، طول

$$S_i = f_i = h_i \cdot I_i$$

$$h_i = \frac{f_i}{I_i}$$

→ با توجه به فراوانی است

$$h_1 = \frac{f_1}{I_1} = \frac{1}{4}$$

$$h_2 = \frac{f_2}{I_2} = \frac{3}{4}$$

$$h_3 = \frac{5}{4}$$

$$h_4 = \frac{2}{4}$$

$$h_5 = \frac{1}{4}$$

نکته: اگر ها فراوانی ها متناسب هم برابر باشند و ... طول دسته ها برابر هم نیستند (I_i ها) طول طبقات

فراوانی نسبی

$$S_i = r_i = h_i I_i$$

$$h_i = \frac{r_i}{I_i} = \frac{r_i}{4}$$

$$h_1 = \frac{r_1}{I_1} = \frac{\frac{1}{12}}{4} = \frac{1}{12\times 4}$$

$$h_2 = \frac{r_2}{I_2} = \frac{\frac{3}{12}}{4} = \frac{3}{12\times 4}$$

$$h_3 = \frac{\frac{5}{12}}{4} = \frac{5}{12\times 4}$$

$$h_4 = \frac{\frac{2}{12}}{4} = \frac{2}{12\times 4}$$

$$h_5 = \frac{\frac{1}{12}}{4} = \frac{1}{12\times 4}$$

کاربردهای هیستوگرام ← ۳

چند درصد داده ها بین ۰-۴ است $\quad \frac{1}{12}\times 100$

چند درصد داده بین ۴-۸ است $\quad \frac{3}{12}\times 100$

چند درصد داده ها کمتر از ۱۲ $\quad \frac{9}{12}\times 100$

چند درصد کمتر از ۲۰ $\quad 100$

چند درصد از داده ها بین ۴ تا ۱۲

چند درصد از داده بین ۶ تا ۹ $\quad \frac{3}{12}+\frac{5}{12}=\frac{8}{12}\times 100$

$$\left((8-6)\frac{3}{12\times 4}+(9-8)\frac{5}{12\times 4}\right)\times 100$$ → تعداد درصد را نشان می دهد

$$\left((8-6)\frac{3}{12\times 4}+(9-8)\frac{5}{12\times 4}\right)\times 12 = (8-6)\frac{3}{4}+(9-8)\frac{5}{4}$$

→ تعداد نفرات را نشان می دهد

چند درصد داده بین داده ۵ تا ۱۷

$$\left((8-5)\times\frac{3}{12\times 4}+\frac{5}{12}+\frac{2}{12}+(17-16)\frac{1}{12\times 4}\right)\times 100$$

Scanned with CamScanner

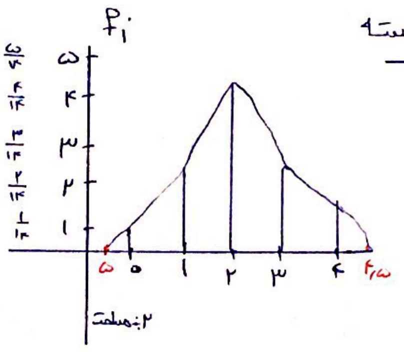
شکل چیز

شکل   چند بر فراوان

چند بر فراوانی نسبی ★ ← به آن تابع چگالی تجربی می‌گویند با افزایش N به سمت تابع چگالی میل می‌کند

چند بر فراوان را به داده‌های گسسته

	f_i	r_i
۰	۱	$\frac{۱}{۱۴}$
۱	۳	$\frac{۲}{۱۴}$
۲	۵	$\frac{۵}{۱۴}$
۳	۳	$\frac{۳}{۱۴}$
۴	۲	$\frac{۲}{۱۴}$
	۱۴	



تابع چگالی چیست ← تابعی است که نحوه‌ی توزیع فراوانی و فراوانی نسبی را (احتمال) روی سطح متغیر تصادفی را به ما نشان می‌دهد

تابع را با $f(x)$ نشان می‌دهیم

$\underset{X}{f}(x)$ ← ۲ شرط را ایجاد می‌کند و ۲ شرط دارد و اگر X با این دو شرط باشد

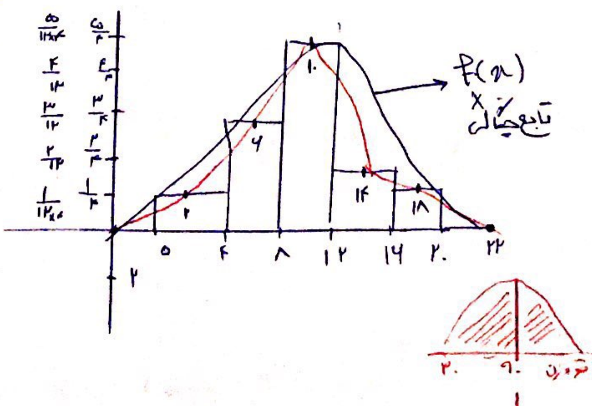
(۱) تابع چگالی زیر محور نمی‌رود و منفی نمی‌شود و همیشه + است ←

$f_X(x) \geq 0$

(۲) مسطح زیر نمودار $f(x)$ یک است

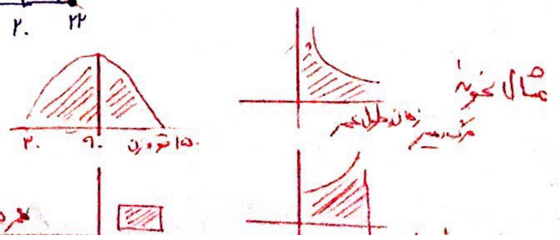
و مساحت زیر آن همیشه یک است

$$\int_{-\infty}^{+\infty} f_X(x)\,dx = 1$$



نکته ← هم $f_X(x)$ است + و هستند نمودار زیر آن + است. سطح زیر نمودار آن + است و باید اینجوری مساحت زیر آن او اهرا است و ۱ را در دارد انتگرال بگیری

مثال نمونه



Scanned with CamScanner

مد

مد ← طبقه‌ای که بیشترین فراوانی را داشته باشد < پیوسته / گسسته

برای داده‌های گسسته

a	f_i
۰	۱
۱	۳
۲	۵
۳	۲
۴	۱

mod = ۲

a	f_i
۰	۴
۱	۳
۲	۵
۳	۵
۴	۱

$mod = \frac{۲+۳}{۲}$

a	f_i
۰	۵
۱	۳
۲	۵
۳	۵
۴	۱

mod = ۰

$mod = \frac{۲+۳}{۲}$

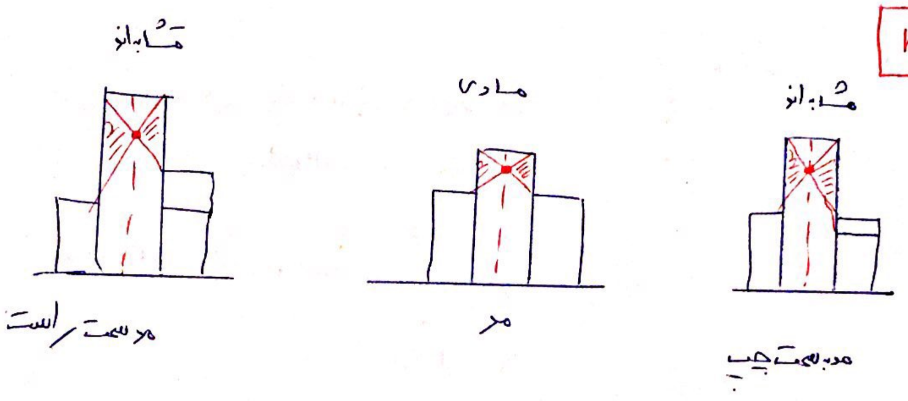
سه مد داریم

a	f_i
۰	۵
۱	۵
۲	۵
۳	۵
۴	۵

مد نداریم

$$mod = L_{i}-1 = \frac{D_1}{D_1+D_2} \cdot I_i$$

برای داده‌های پیوسته



$$D_1 = f_i - f_{i-1}$$

$$D_2 = f_i - f_{i+1}$$

Scanned with CamScanner

$$\bar{x} = \frac{1}{\bar{x}} = \frac{1}{n}\sum f_i \frac{1}{x_i}$$

$$H = \frac{n}{\sum \frac{f_i}{x_i}}$$

میانگین هارمونیک

مثال ← میانگین سرعت؟؟

با سرعت ۵۰ رفتم قم و با سرعت ۶۰ برگشتم. متوسط سرعتم چقدر؟؟

$$\xleftarrow[v_2]{\text{سرعت}}$$

تهران ——— d ——— قم

$$\xrightarrow[\text{سرعت}]{v_1}$$

مجموع مسافت / مجموع زمان‌ها

فرمول $\bar{V} = \frac{\sum d_i}{\sum t_i}$ ← مجموع مسافت / مجموع زمان

$V = \frac{d}{T}$ ، $T = \frac{d}{V}$

$$= \frac{\sum d_i}{\sum \frac{d_i}{v_i}} = \frac{d_1 + d_2}{\frac{d_1}{v_1} + \frac{d_2}{v_2}}$$

$$= \frac{2d}{d\left(\frac{1}{v_1} + \frac{1}{v_2}\right)} = \frac{2}{\frac{1}{v_1} + \frac{1}{v_2}}$$

میانگین هندسی

$$G = \sqrt[n]{x_1^{f_1} \cdot x_2^{f_2} \cdots x_k^{f_k}}$$

$$G = \sqrt[n]{\prod_{i=1}^{k} x_i^{f_i}}$$

مثال = نرخ تورم امسال ۱۲٪ و سال بعد ۱۵٪ و سال بعدش ۲۰٪

متوسط نرخ تورم در این ۳ سال چقدر است

* برای نرخ می‌توانیم از میانگین هندسی استفاده کنیم به دلیل

$$A_1 = 12 \times A$$

$$A_2 = 15 \times A_1 = 15 \times 12 \times A$$

$$A_3 = 20 \times A_2 = 20 \times 15 \times 12 \times A$$

A

$A_1 = x \cdot A$

$A_2 = x A_1$

$A_3 = x A_2 = x^3 A$

$20 \times 15 \times 12 \times A = x^3 A$

$$x = \sqrt[3]{20 \times 15 \times 12}$$

* مهم

$$H \leq G \leq \bar{x} \leq \sqrt{\frac{\sum x^2}{n}}$$

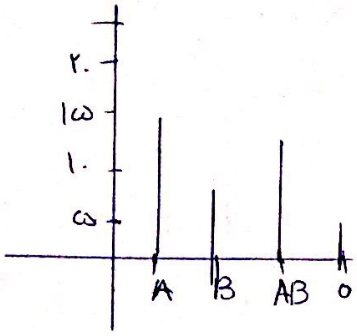
هارمونیک — هندسی — حسابی — مجذور میانگین مربعات

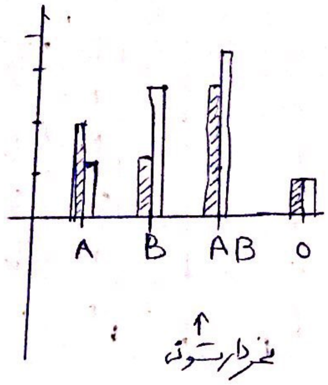
$$\sqrt[n]{1 \times 2 \cdots \times n} \leq \frac{\sum_{i=1}^{n}}{n}$$

آمار توصیفی

نمودار میله‌ای (ستونی)

	Rh^+	Rh^-	
A	۱۵	۱۰	۲۵
B	۱۰	۱۵	۲۵
AB	۱۵	۲۰	۳۵
O	۵	۵	۱۰
			۹۵





↑
نمودار ستونی

* مرکز ۲ صفت دارد در جامعه جایگاه دارد اگر داده‌ها را در عدد n ضرب کنیم و با یک عدد جمع کنیم و شاخص‌ها در آن عدد ضرب و در آن عدد جمع می‌شوند

شاخص‌ها < مرکز (مکانی) < میانگین < هندسی / حسابی ؛ میانه ؛ مد
پراکندگی

میانگین حسابی ← مجموع انحراف داده‌ها از میانگین حسابی همواره برابر صفر است

خصوصیات ۳ مورد است

① $\bar{X} = \frac{1}{n}\sum_{i=1}^{k} (f_i)(x_i) = \sum_{i=1}^{k} \frac{f_i}{n} x_i = \sum_{i=1}^{k} r_i x_i$

(فراوانی / تعداد داده؛ در نمونه؛ در جامعه $\sum p_i x_i$)

مثال جدول در ترم آخر

امید ریاضی

مثال ساده‌تر:

۱۲۴ → $\bar{x} = ۳$ میانگین — انحراف: $-۲ \ -۱ \ +۳ = ۰$

۱۲۳ → $\bar{x} = ۲$ میانگین — انحراف: $-۱ \ ۰ \ ۱ = ۰$

$$\sum_{i=1}^{k} f_i (x_i - \bar{x}) = 0$$

(مجموع انحراف داده از میانگین برابر صفر است)

② $y_i = a x_i + b$

آموزش به صورت ریاضی

میانگین $\bar{y} = \frac{1}{n}\sum f_i y_i$

$= \frac{1}{n}\sum f_i (a x_i + b)$

$= \frac{1}{n}\sum f_i a x + \frac{1}{n}\sum f_i b$

$= a \frac{1}{n}\sum f_i x_i + b \frac{1}{n} \sum f_i = a\bar{x} + b$

میانگین توابع داده

③ $\overline{h(x)}$

$h(\bar{x}) = \frac{1}{n}\sum f_i h(x)$

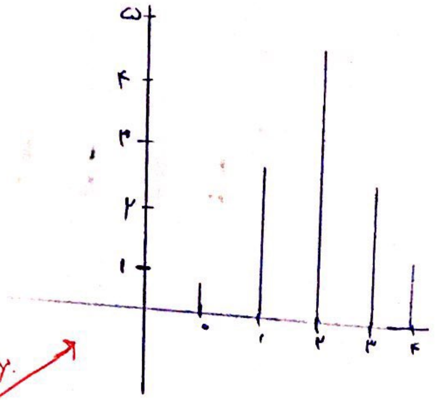
$\bar{x^2} = \frac{1}{2}\sum f_i x_i^2 \rightarrow \sqrt{\bar{x^2}}$

$\bar{x^3} = \frac{1}{n}\sum f_i x_i^3 \rightarrow \sqrt[3]{\bar{x^3}}$

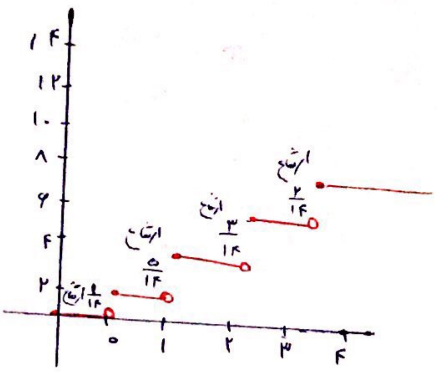
Scanned with CamScanner

X	f_i	r_i	F_i	R_i
۰	۱	$\frac{1}{14}$	۱	$\frac{1}{14}$
۱	۳	$\frac{3}{14}$	۴	$\frac{4}{14}$
۲	۵	$\frac{5}{14}$	۹	$\frac{9}{14}$
۳	۳	$\frac{3}{14}$	۱۲	$\frac{12}{14}$
۴	۱	$\frac{1}{14}$	۱۴	۱
	۱۴	۱		

نمودار فراوانی تجمعی (فراوانی تجمعی نسبی) (تابع توزیع تجربی)



بر اساس f_i (فراوانی)



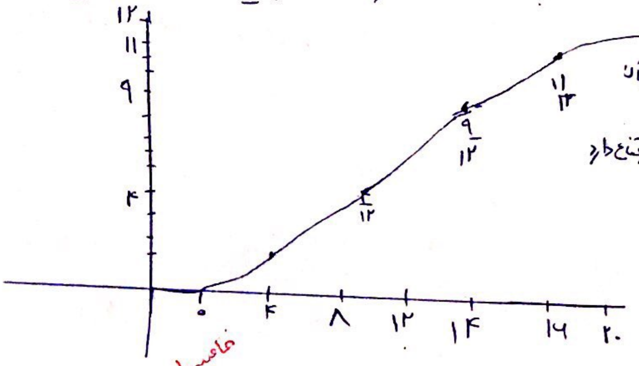
نمودار فراوانی تجمعی

در داده ها گسسته (وزن روی نقطه است)

پیوسته

* نکته ← تابعی است که از راست پیوسته و حد اقل ۰ و حد اکثر ۱ می باشد

X	f_i	r_i	F_i	R_i
۰ - ۴	۱	$\frac{1}{12}$	۱	$\frac{1}{12}$
۴ - ۶	۳	$\frac{3}{12}$	۴	$\frac{4}{12}$
۶ - ۸	۵	$\frac{5}{12}$	۹	$\frac{9}{12}$
۸ - ۱۰	۲	$\frac{2}{12}$	۱۱	$\frac{11}{12}$
۱۰ - ۱۲	۱	$\frac{1}{12}$	۱۲	۱
	۱۲	۱		



فاصله طبقات ۴ تا ۴ میباشد

* نمودار فراوانی تجمعی نسبی ← را تابع توزیع می گویند با افزایش N تعداد نمونه تابع توزیع به یک تابع میل می کند و که به آن تابع تابع توزیع گویند $f_X(x)$ نشان می دهیم

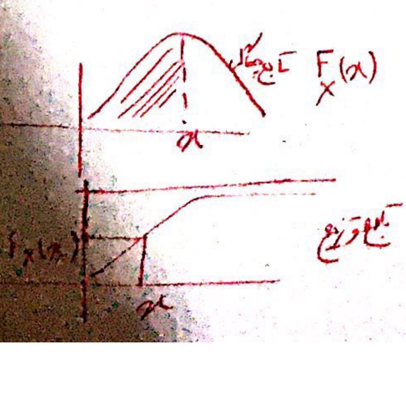
* تابعی است صعودی (غیر نزولی) آن کمتر و کوچکتر از یک است حد اقل از راست پیوسته است

نکات، خصوصیات تابع

1) $F_X(-\infty)=0 \quad F_X(+\infty)=1$

2) $x_1<x_2 \to F(x_1) \le F_X(x_2)$

3) $\lim_{x \to a} F_X(x) = F_X(a)$



رابطه بین تابع توزیع و تابع چگالی (رابطه آن مشتق و انتگرال است. و چیزی جز بدست آوردن مساحت هانیست)

$$F_X(x)' = f_X(x)$$

$$\int_{-\infty}^{x} f_X(x)\, dx = F_X(x)$$

Scanned with CamScanner

جلسه سوم — مهر ۷

۳ پارامتر ← ضریب تغییرات ← C.V $C.V = \frac{S}{\bar{x}} \times 100$
← ضریب چولگی
← ضریب کشیدگی

ضریب چولگی (چولگی یعنی عدم تقارن) اگر توزیع متقارن نباشد به سمت دیگر متمایل می‌شود
اگر فرض داشته باشیم توزیع را توزیع چوله می‌نامیم

ضریب چولگی

۳ ضریب داریم (چولگی)

$$SK_1 = S_{k_1} = \frac{\bar{x} - mod}{s}$$

$$SK_2 = S_{k_2} \quad \frac{3(\bar{x} - med)}{s}$$

$$SK_3 = \frac{m^3}{s^3} = \frac{m^3}{(\sqrt{s^2})^3}$$ (برای سنجش کشیدگی داده‌ها مهم)

اول باید گشتاور مرکزی را بدست آوریم و جذر بگیریم و بعد به توان ۳ برسانیم

واریانس ← اول جذر می‌گیریم بعد به توان ۳ برسانیم

در حالت کلی اگر sk اگر از $\frac{1}{10}$ یک دهم کوچکتر باشد و از $\frac{1}{10}$ بزرگتر باشد می‌پذیریم این نمونه‌ای که گرفتیم از جامعه متقارن آمده است

نکته: اگر sk بین ۰/۱ و ۰ بود تا بزرگتر و مساوی با ۰/۱ بود می‌گوییم چوله به راست

در حالت کلی ← $-0/1 < S_k < 0/1$

یا ← $|S_k| < 0/1$

ولی اگر کوچکتر و مساوی با ۰/۱- بود می‌گوییم چوله به چپ

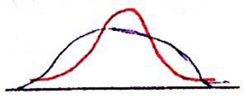
و می‌گوییم در این حالت شاخص داریم

مفهوم کشیدگی =

آموختیم که پراکندگی بد بود چون باعث ریسک می‌شود

مثلاً ما یک سری داده روی خط داریم و پخش شده است و حداکثر میزان پراکندگی آنها است و اگر به داده‌ها از دو طرف فشار بیاوریم این داده‌ها شروع می‌کنند جمع شدن و داده‌ها فشرده می‌شوند روی هم می‌آیند و پراکندگی آنها کم‌تر می‌شود جمع می‌شوند حول معیار مرکزی و قله‌ی جامعه به سمت بالا می‌رود و کشیدگی اتفاق می‌افتد هرچه از بغل‌ها بازتر بشیم پخش اتفاق می‌افتد و قله می‌آید پایین و داده‌ها باز می‌شود و به آن می‌گویند داده‌ها پخش شده‌اند.

مثال دانه‌های تسبیح:



از بالا که داده‌ها می‌ریزند وسط تمرکز زیاده و هرچه گوشه‌تر می‌رویم تمرکز کم‌تر می‌شود و اگر چسبندگی داشته باشند اگر داده‌ها را از بالا بریزند اینها روی هم جمع می‌شوند و کم‌تر پخش می‌شوند یعنی کشیدگی داده‌ها زیاد می‌شود

هرچه کشیدگی بیشتر تمرکز حول مرکز بیشتر
هرچه کشیدگی کمتر تمرکز حول مرکز کمتر

کشیدگی زیاد یک معیار خوبی هست چون کشیدگی باعث می‌شود پراکندگی کم شود

Scanned with CamScanner

ما دو تا گشتاور داریم ← ۱) مرکزی ۲) غیر مرکزی

* مرکزی

$$m_r(a) = \frac{1}{n}\sum f_i (x_i - \bar{x})^r$$

$$m_1 = \frac{1}{n}\sum f_i (x_i - \bar{x}) = 0$$

مجموع انحرافات داده‌ها از میانگین حساب برابر صفر است

$$m_2 = \frac{1}{n}\sum f_i (x_i - \bar{x})^2 = S^2$$

یا $= \bar{x^2} - \bar{x}^2 = m'_2 - (m'_1)^2 = S^2$

$$m_3 = \frac{1}{n}\sum f_i (x_i - \bar{x})^3 = \bar{x}^3$$

* غیر مرکزی (زیاد هستن و فقط صفر مهم است)

$$m'_r = \frac{1}{n}\sum f_i x_i^r$$

$$m'_1 = \frac{1}{n}\sum f_i x_i^1 = \bar{x}$$

$$m'_2 = \frac{1}{n}\sum f_i x_i^2 = \bar{x}^2$$

$$m'_3 = \frac{1}{n}\sum f_i x_i^3 = \bar{x}^3$$

خواص گشتاور ←

گشتاور غیر مرکزی ← اگر داده‌ها را در عدد ضرب و یا به عدد جمع کنیم گشتاور چه تغییری می‌کند ؟؟؟

غیر مرکزی = گشتاور غیر مرکزی فقط m'_1 است که همان عدد ضرب و با همان عدد جمع می‌شود و در مورد بقیه صدق نمی‌کند

گشتاور مرکزی = ۱) اگر داده‌ها در یک عدد ضرب و یا به عدد جمع کنیم فقط گشتاور مرکزی فقط در a^r ضرب می‌شود

۲) تمام گشتاورهای مرکزی می‌توانیم بر اساس گشتاور غیر مرکزی هم رتبه می‌توانیم بنویسیم

* ۳) در جوامع متقارن گشتاورهای مرکزی رتبه فرد همواره برابر صفر است .

نکته ** مجموع انحرافات داده‌ها از میانگین حسابی برابر با صفر است

Scanned with CamScanner

هشت

مفهوم درجه آزادی

مثال ← ۲ درجه آزادی دارد $d.f = 2$ $\quad X_1 \quad X_2$ ما دوتا متغیر داریم ← ما دوتا متغیر داریم به نام x_1 , x_2

x ها همه فرضی هستند و x_1 , x_2 هر مقدار
و هر عددی را می‌توانند بگیرند ولی x_1 , x_2 هر مقدار
در ... می‌توانند داشته باشند

مثال ← ۱ درجه آزادی دارد $d.f = 1$ $\quad \frac{x_1 + x_2}{2} = 150$

مثال ← درجه آزادی ندارد $d.f = 0$ $\quad x_1 - x_2 = 15$ دو تا معادله دو تا مجهول داریم

وقتی درجه آزادی بالا باشد روی داده‌ها می‌توانیم راحت‌تر کار کنیم - وقتی درجه آزادی زیاد باشد کاری که انجام می‌دهیم دقتش بیشتر است)

برای همین در جوامع نمونه را زیاد می‌کنیم تا درجه آزادی را بالا ببریم (با افزایش درجه آزادی بالا بالا دقت بیشتر است)

سوال ← ۲ تا فرمول برای واریانس تعریف می‌کنیم

به نظر شما کدام یک از این دو فرمول دارای دقت بیشتری است چرا؟؟

$$S_u^2 = \frac{1}{n-1} \sum f_i (x_i - \bar{x})^2$$

مهم‌تر است — واریانس نااریب

$$S_b^2 = \frac{1}{n} \sum f_i (x_i - \bar{x})^2$$

واریانس اوریب یا کجی

جواب: ۱- اولی دقت بیشتری دارد

ما در محاسبه واریانس از میانگین استفاده می‌کنیم

ولی از داده‌ها دیگر اهمیت ندارد و باید تقسیم بر $\frac{1}{n-1}$ کنیم چون میانگین را از روی داده‌ها به دست می‌آوریم

۲- دقت آن بالاست ولی محاسبه آن سخت‌تر است

★ واریانس نااریب مهم‌تر است چون در محاسبه واریانس از میانگین استفاده می‌کرد، $\frac{1}{n-1}$

سوال ← اگر $n \to \infty$ اگر n به سمت بی‌نهایت میل کند چه اتفاقی می‌افتد

$n \to \infty$
$S_u^2 \approx S_b^2$

در آمار حد ۳۰ یا ۲۵ است

نکته — اگر تعداد نمونه بیشتر از ۲۵ یا ۳۰ باشد فرق نمی‌کند از کدام S_b^2 - S_u^2 استفاده کنیم

نکته — اگر تعداد نمونه کمتر از ۲۵ بود باید از S_u استفاده کنم

فرمول

$$S_V^2 = \frac{n}{n-1} \quad S_b^2 \frac{n}{n-1} (\bar{x}^2 - \bar{x}^2) = \frac{1}{n-1}$$

$$\left(\sum x_i^2 - \left(\frac{\sum x_i}{n} \right)^2 \right)$$

۱ مجموع داده‌ها به توان ۲

Scanned with CamScanner

$R = \max(x_i) - \min(x_i)$

$D = \sum_i |x_i - \bar{x}|$

$AD = \frac{1}{n} \sum f_i |x_i - \bar{x}|$

واریانس $s^2 = \frac{1}{n} \sum f_i (x_i - \bar{x})^2$

جذر می‌گیریم

انحراف معیار $S = \sqrt{s^2}$
انحراف مطلق دارد

ضریب تغییرات $CV = \frac{s}{\bar{x}}$ معیار

معیار مرکزی

هرچه کوچکتر باشد داده‌ها همگن‌تر هستند
هرچه بزرگتر باشد پراکندگی بیشتر است

تابع چگالی و تابع توزیع ↔ ارتباط این دو تابع مشتق و انتگرال هست

۲ روش داریم برای مثبت کردن و برای اینکه عدد منفی نشود:
- از قدر مطلق استفاده کردن
- یا به توان ۲ رساندن

مثلاً ← یک جامعه داریم قد زرافه‌های بالغ در دنیا.

R دامنه تغییرات چقدر است؟

بین ۴ متر تا ۶ متر ← اختلاف ۲

مثلاً ← جامعه قد بالغ مورچه‌ها در دنیا ←

R دامنه تغییرات چقدر

که کوچک‌ترین مورچه اصلی قد و بزرگترین. اصلی متر (میلیمتر) (اندازه‌اش)

(پراکندگی قد در جامعه مورچه‌ها از جامعه زرافه‌ها بیشتر است)

فرمول واریانس $s^2 = \frac{1}{n} \sum f_i (x_i - \bar{x})^2$

$s^2 \geq 0$

* واریانس چه خواصی دارد
(اگر داده‌ها را در عددی ضرب یا عددی جمع کنیم مثلاً چه تأثیری در آن عدد ضرب می‌شود)

$s^2 \geq 0$ بزرگتر از صفر است و همواره + است

اگر داده‌ها را در a ضرب و با b جمع $\quad y_i = a x_i + b$

$$s^2_y = \frac{1}{n} \sum_{i=1}^{k} f_i (y_i - \bar{y})^2$$

اگر داده‌ها را با عددی جمع کنیم هیچ تأثیری در واریانس ندارد

$$= \frac{1}{n} \sum f_i \left(a x_i + b - (a\bar{x} + b)\right)^2$$

$$= a^2 \frac{1}{n} \sum f_i (x_i - \bar{x})^2 = a^2 s_x^2$$

Scanned with CamScanner

مفاهیم پراکندگی ←

$$R = \max(x_i) - \min(x_i)$$

x	f_i	F_i
0	50	$50\,\lvert 100-50\rvert = 2500$
100	50	$50\,\lvert 100-50\rvert = 2500$
		5000

$\bar{x} = 50 = med$

$\bar{x} = 50\ med$ دامنه تغییرات

$R = 100 - 0 = 100$

اختلاف ۱۰۰ واحد است

ریسک جامعه ۱ ← ریسک زیاد است و پراکندگی زیاد است ← و R قابل اعتماد نیست

x	f_i	F_i
49	50	$50\,\lvert 49-50\rvert = 50$
51	50	$50\,\lvert 51-50\rvert = 50$
		100

$x = 50\ med$

$R = 51 - 49 = 2$

اختلاف ۲ واحد است

ریسک جامعه ۲ ← ریسک کمتر است ←

نکته) R ** R خیلی ساده به دست می‌آید ولی قابل اعتماد نیست زیرا دو تا داده اثر دارد کوچک‌ترین و بزرگ‌ترین

x	f_i	$F_i\lvert x_i - 50\rvert$
0	1	1×50
49	49	49×1
51	49	49×1
100	1	1×50
	100	198

$\bar{X} = med = mod = 50$

$R = 100$

مثال)

x	f_i	$F_i\lvert x_i - 50\rvert$
0	1	1×50
100	1	1×50
	2	100

$\frac{100}{2} = 50$

به نسبت افراد جامعه تقسیم می‌کنیم

* تعداد افراد نمرات برابر $\bar{x} = 50 = med$ در میزان پراکندگی اثر ندارد

جامعه اول است ولی اعداد جامعه دومی هستند

فرمول $D = \sum_i \lvert x_i - \bar{x}\rvert$ ← مجموع فواصل داده‌ها از میانگین

نکته) هرچه مجموع فواصل بیشتر پراکندگی بیشتر است

هرچه مجموع فواصل کمتر پراکندگی کمتر است

نکته) D هم معیار خوبی نیست زیرا از n به دلیل مثال روبرو است

مثال ← معدل کلاس ۱۰ است و پراکندگی کم است یعنی چه ← یعنی بچه‌ها بین ۸ تا ۱۲ هستند و حول مرکز دارند تغییر می‌کنند

مثال ← یک جامعه را مثال بزنید که پراکندگی صفر باشد ← زمانی که آخر ترم امتحان گرفته شود و همه ۲۰ شوند و اگر یک نفر ۹ شود باعث ایجاد پراکندگی می‌شود

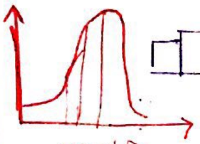
Scanned with CamScanner

معیارهای مرکزی را توضیح داده شده معیارهای مرکزی توی داده ها دارای مکان هستند

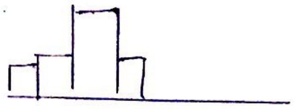
تقارن ← اگر از یکی معیارهای مرکزی به دلخواه (میانگین - میانه - مد) یکی را انتخاب کنیم و از آن به فاصله های یکسان به دو طرف حرکت کنیم فراوانی یا فراوانی نسبی یکسان می بینیم

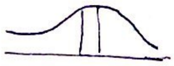
mod
x

- در جامعه متقارن یا (متقارن) میانگین، میانه، مد با هم برابراند
مربع جهت تغییرات مد
بزرگترین تغییرات در میانگین ایجاد می شود

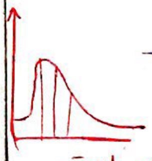


چوله به چپ

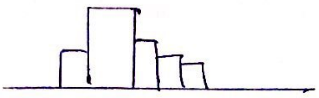




- در جامعه چوله به چپ $\bar{x} < med < mod$



چوله به راست



- در جامعه چوله به راست $mod < med < \bar{x}$

در جامعه متقارن ←

$$\bar{x} - mod \approx ۳(\bar{x} - med)$$

اگر راست چوله +
اگر چپ چوله -

(رابطه ی پیرسون) اگر رابطه تقریباً برابر باشد می پذیریم نمونه ای که گرفتیم از جامعه متقارن

در داده ها دارای مکان هستند
اگر داده ها را در عددی ضرب یا با یک عدد جمع کنیم (آنها) را در آن ضرب و در آن عدد جمع می کنیم
میانگین - مد - میانه

معیارهای مرکزی
معیارهای پراکندگی

معیارهای پراکندگی

دارای مکان نیستند
پراکندگی صفرش مطلق است یا هست یا وجود دارد یا وجود ندارد
اگر داده ها را در عددی ضرب یا با یک عدد جمع کنیم فقط معیارهای پراکندگی در آن عدد ضرب می شود و + و - روی میزان پراکندگی اثری ندارد
یعنی (مقایسه جامعه مهم است مکان مهم نیست)

* تقسیم بندی در جامعه ای که دارای پراکندگی کمتر است آسانتر است

+ پراکندگی در تقسیم بندی راحت نیست

پراکندگی بیشتر است و تمرکز کم است

پراکندگی کمتر است
تمرکز بیشتر است

$\sqrt{\bar{x}^2} < \bar{x} =$ یا مقدار میانگین رتبه ۲

Scanned with CamScanner

میانه برای داده‌های پیوسته

$$mod = \frac{L_i + \frac{n}{2} + F_{i-1}}{2}$$

$$= L_3 + \frac{6 - F_2}{f_3} \cdot I_3$$

حد پایین طبقه سوم — فراوانی همان طبقه f_3 — طول طبقه

x	f_i	F_i (فراوانی تجمعی)
۰ – ۴	۱	۱
۴ – ۸	۳	۴
★ ۸ – ۱۲	۵	۹
۱۲ – ۱۶	۲	۱۱
۱۶ – ۲۰	۱	۱۲
	$n = 12$ (مجموع)	

$$= 8 + \frac{6 - 4}{5} \times 4 = 9.6$$

حد پایین — فراوانی تجمعی ماقبل — فراوانی همان طبقه — طول طبقه

اول باید $\frac{n}{2}$ را بدست آوریم $n = \frac{12}{2} = 6$

$i = 3$ و جوابی که بدست می‌آوریم می‌تواند از ۸ کوچک‌تر و از ۱۲ بزرگتر باشد

میانه برای داده‌های گسسته

$$mod \begin{cases} \dfrac{x(\frac{n}{2}) + x(\frac{n}{2}+1)}{2} & \text{زوج} \\[2ex] x(\frac{n+1}{2}) & \text{فرد} \end{cases}$$

x	۰	۱	۲	۳	
f_i	۲	۴	۳	۳	۱۲
F_i (فراوانی تجمعی)	۲	۶	۹	۱۲	

$n = 12$ مجموع — زوج هست

۰۰ ۱۱۱۱ ۲۲۲ ۳۳۳

$$x(6) + x(7) = \frac{1+2}{2} = 1.5$$

- میانه تحت تأثیر داده‌های بزرگ و کوچک نیست و فقط تعداد افراد مهم است
- میانه از نظر فیزیکی کجاست ؟؟ از این داده‌های داده‌های پیوسته باشد دقیقاً میانه جایی است که سطح زیر نمودار را نصف می‌کند
- میانه عددی است که ۵۰٪ داده‌ها بعد از آن هستند و ۵۰٪ آن قبل از آن هستند

کاربرد مهم میانه کجاست؟

$$\text{کاربرد میانه} \rightarrow \underset{\text{مجموع}}{\sum} \underset{\text{خاصه}}{f_i} \, |\underset{\text{داده‌ها}}{x_i} - mod| \leqslant \sum f_i \, |x_i - a|$$

← فرمول کاربرد میانه

به فارسی: مجموع فواصل مطلق داده‌ها از میانه > از مجموع فواصل داده‌ها از هر عدد دیگری کمتر است

Q_p (کوانتایل‌ها) عددی است $100 \times p$ $\quad 0 < p < 1$

مفهوم کلی میانه ← چندک‌ها (کوانتایل‌ها)، دهک‌ها، صدک‌ها

– میانه حالت خاصی از کوانتایل‌ها است و چند کوانتایل‌های معروف داریم عبارت Q_1 Q_2 Q_3 که به آنها چارک می‌گویند

Q_1	Q_2	Q_3
عددی است که ۲۵٪ داده‌ها قبل از آن می‌آیند	همان میانه است عددی که ۵۰٪ داده‌ها که قبل از آن می‌آیند	عددی که ۷۵٪ از داده‌ها قبل از آن می‌آیند

* این ۳ نقطه با لا جامعه را به ۴ دسته تقسیم کرده به ۴ تا ۲۵٪ تقسیم کرده است تعداد افراد مهم است

دهک‌ها ← $D_1 \; D_5 \; \cdots \; D_9$

صدک‌ها ← $P_1 \; P_9 \; \cdots \; P_{99}$

+ مثلاً یارانه‌ها را به طوری سیستم کنیم که ۳ دهک پایین جامعه بیشترین کمک داشته باشد؟ اگر بگوییم که دهک سوم جامعه از نظر حقوق برابر با ۲ میلیون تومان است یعنی ۳۰٪ افراد جامعه حقوق کمتر از ۲ میلیون دارد

+ یا دهک ۸ جامعه از نظر حقوق ۵ میلیون است یعنی ۸۰٪ جامعه از حقوق کمتر از ۵ میلیون تومان می‌گیرند

Scanned with CamScanner

داده ها گسسته

فرمول

$$Q_p \begin{cases} x_r & (n+1)p = r \text{ صحیح} \\ x_{(r)} + \omega (x_{(r+1)} - x_{(r)}) & (n+1)p = r,\omega \text{ اعشاری} \end{cases}$$

مثال ← مثال Q_1 را حساب کنید و بعد Q_3 را حساب کنید

$Q_1 = Q_{0.25}$

$(12+1) \times 0.25 = 3.25$ اعشاری

$r = 3 \quad \omega = 0.25$

فرمول ← $Q_1 = Q_{0.25} = x_{(3)} + 0.25(x_{(4)} - x_{(3)})$

$= 1 + 0.25(1-1) = 1$

عددی است که ۲۵٪ داده ها را قبل از آن باشد و ۷۵٪ داده ها بعد از آن

* میانه حالت خاصی از کوانتایل ها می باشد

* کوانتایل ها مانند میانه عمل می شود

مثال ← D_8 را حساب کنید

$$D_8 = Q_{0.8} = x_{(10)} + 0.3(x_{(11)} - x_{(10)}) = 3 + 0.3(3-3) = 3$$

عدد ۳ عددی است که می توان ۸۰٪ داده ها را قبل از آن قرار داد

داده های پیوسته

$$Q_p = L_i + \frac{n.p - F_{i-1}}{f_i} . I_i$$

	f_i	F_i
۰ – ۴	۱	۱
۴ – ۸	۳	۴
۸ – ۱۲	۵	۹
۱۲ – ۱۶	۲	۱۱
۱۶ – ۲۰	۱	۱۲
	$n=12$	

اولین طبقه ای که فراوانی تجمعی آن برابر یا بیشتر از np می باشد

مثال ← D_7 را حساب کنید

$D_7 = Q_{0.7}$

$12 \times 0.7 = 8.4$

$$D_7 = Q_{0.7} = L_3 + \frac{8.4 - F_2}{F_3} . I_3$$

$$12 \times 0.7 = 8.4 = \frac{8 + 8.4 - 4}{5} \times 4$$

$i = 3$

در طبقه دوم قرار می گیرد

$= 11.52$ از ۱۲ کمتر است

۷۰ درصد داده ها قبل از آن است

مثال ← Q_3 را حساب کنید

Q_3 یا ۷۵٪

$12 \times \frac{3}{4} = 9$

$i = 3$

$$Q_3 = L_3 + \frac{9 - F_2}{F_3} . I_3 = 12$$

عددی است که ۷۵٪ داده ها قبل از آن قرار می گیرد

$12 \times \frac{3}{4} = 9$

Scanned with CamScanner

مثال ادامه ←

مالکشیدگی را با چی می‌سنجیم؟ در تقارن تعریف مشخص بود جامعه متقارن نه چولگی چپ داشت نه چولگی راست

اما در کشیدگی را به عنوان ملاک قرار دهیم از ارتفاع بالاتر کشیده می‌شود و از ارتفاع کم تر پهن می‌شوند پخ کجا؟؟

دلیل برای فهمیدن آن نیاز داریم به بحث دیگری شویم

تابع چگالی نرمال ↓

تابع ⟨ توزیع / چگالی — که قبلاً خوانده ایم تابع چگالی تابعی است مثبت که سطح زیر نمودار آن همواره برابر با یک است

تابع چگالی متفاوتی وجود دارد یعنی شکل ها فرق دارند معروف ترین نوع چگالی این تابع است که در آن

و یک تابع نمایی است و شکل تقریبی آن توزیع نرمال مشهور

$$f_x(x) = \frac{1}{\sqrt{2\pi s}} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2}$$

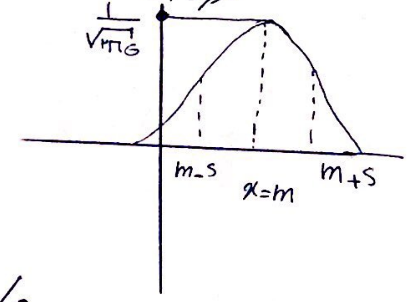
سیگما

اعداد حقیقی
$x \in R$ یعنی x می‌تواند هر عددی باشد

$m \in R$ m می‌تواند هر عددی باشد

$s > 0$ سیگما باید مثبت

π , e
π → ۳٫۱۴ ، e → ۲٫۷۸



خصوصیات
نکات مهم برای رسم شکل

۱- (۱) محور x ها مجانب افقی

۲- (۲) $x = m$ محور تقارن

۳- (۳) max x = x = ، $y = \frac{1}{\sqrt{2\pi\sigma}}$

۴- (۴) $x = m \pm s\,\sigma$ سیگما

شکل چگالی نرمال

شکل تقریبی آن

⟨ملاک تشخیص کشیدگی توزیع نرمال برازش شده داده هاست⟩ **

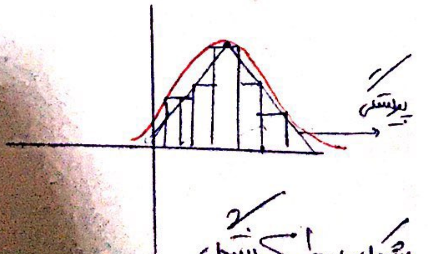
برازش داده هایعنی ← ربط به داده ها می‌شود ممکن است داده ها نرمال باشد و یا ممکن است داده ها نرمال نباشد

و برازش داده ها بر اساس

میانگین
$m = \bar{x} = \frac{1}{n} \sum f_i x_i$

$\sigma^2 \simeq s^2 = \frac{1}{n} \sum f (x_i \cdot \bar{x})^2$

سیگما، اس دو



شکل مربوط کشیدگی

برای تشخیص کشیدگی می‌توانیم هیستوگرام داده ها را کشیم یا رسم کنیم

و کاری نداریم که داده ها چی هستند اول میانگین را از روی داده ها حساب می‌کنیم و بعد واریانس را

حساب می‌کنیم و در می‌گذاریم داخل معادله و می‌دهیم به هیستوگرام رسم می‌کنیم که از میانگین و واریانس

این جامعه بدست می‌آید و بعد نمودار یا چندضلعی فراوانی نسبی را با آن داده رسم می‌کنیم اگر چند ضلعی بزرگتر بالاتر جامعه کشیده می‌شود و از پایین تر پهن می‌شود ←

Scanned with CamScanner

← ملاک تشخیص آن مربوطش k هست

$$K = \frac{m_4}{s^4} - 3$$

$|K| < 0.1$

اگر |k| کوچکتر از ۰٫۱ باشد می‌پذیریم که این جامعه نسبت به توزیع نرمال نه کشیده است نه پخ است

$|K| < 0.1$

اگر بیشتر از ۰٫۱ باشد کشیده هست

اگر کمتر از ۰٫۱- باشد پخ هست

** کشیدگی خوب می‌باشد چون پراکندگی داده‌ها حول مرکز کم است

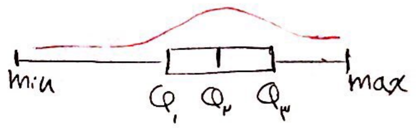
معرفی ← جداول ناقص و جداول کامل

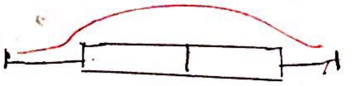
جدول ناقص است

$x < 10$	چون ابتدای طبقه اول مشخص نیست و نماینده ندارد
۱۰_۲۰	
۳۰_۴۰	
$x > 40$	انتهای طبقه آخر مشخص نیست و نماینده ندارد.

میانگین و هر چولگی ... محاسبه نمی‌شود و ناقص است

نمودار جعبه‌ای یا باکس پلات ← با عمودی یا افقی رسم می‌شود و می‌تواند شکل‌های مختلفی باشد تقارن یا [illegible] داشته باشد





★ ← فرق آنها پراکندگی در آن بیشتر است و پخ‌تر است

تمرکز

تمرکز کم است

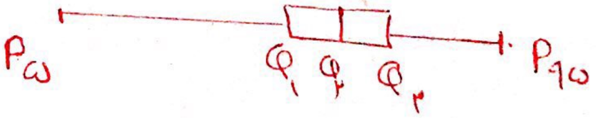
چوله به چپ

چوله راست — تمرکز داده در این نقطه زیاد است

تمرکز زیاد است

* داده‌های پرت: داده‌هایی هستند که خیلی بزرگ‌اند یا خیلی کوچک‌اند و مشکلاتی را ایجاد می‌کنند و میانگین را به سمت خودشان می‌کشند

برای اینکه داده‌ها پرت (پرت) اثر خود را از دست بدهند ۵٪ از ابتدا و ۵٪ از انتها حذف می‌کنیم — داده‌ها پرت الزاماً بد نیستند و بعضی داده‌ها پرت را حذف نکنیم داده از خوش اندازه می‌کنیم و میانگین به دست آمده میانگین پیراسته و واریانس پیراسته



← اگر ۳ تا برابر و برقرار باشد می‌توانیم بگوییم جامعه متقارنه (مفهوم تقارن در باکس پلات)

$$\frac{max + min}{2} = Q_2$$

$$Q_3 - Q_2 = Q_2 - Q_1$$

$$\frac{Q_1 + Q_3}{2} = Q_2$$

ادامه

Scanned with CamScanner

← ادامه

هرچه کمتر پراکندگی بیشتر
هرچه بیشتر پراکندگی کمتر

مرکزی $\frac{Q_3 - Q_1}{2} \simeq Q_2$

پراکندگی $Q_3 - Q_1$

$\frac{\text{ضریب تغییرات}}{\text{میانه}} = b = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_2) + (Q_2 - Q_1)}$

مثال ←

دو تا جامعه داریم n_1	$\bar{x}_1$	s_1^2
n_2	$\bar{x}_2$	s_2^2
⋮	⋮	⋮

$n_1 + n_2$

$$\bar{x} = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2}{n_1 + n_2}$$

(واریانس آمیخته) $$S_p = \frac{n_1 s_1^2 + n_2 s_2^2}{} + \frac{n_1(\bar{x} - \bar{x}_1)^2 + n_2(\bar{x} - \bar{x}_2)^2}{n_1 + n_2}$$

اصل ترکیب < اشیا

اصل ترکیب < اشیاء متمایز هستند / تکرار مجاز نیست

ترتیب

اگر بگوید ترتیب مهم نیست (جابه‌جایی مهم نیست)

ترتیب: ۱) اشیاء متمایزند ۲) تکرار مجاز نیست

$P^n_n = n.(n-1) \cdots \times 1 = n!$

$P^n_r = n(n-1)(n-2)\cdots(n-r+1) \times \frac{(n-r)!}{(n-r)!} = \frac{n!}{(n-r)!}$

مثال ← n نفر می‌خواهند در یک صف n تایی قرار بگیرند. چند طریق می‌توانند در صف قرار بگیرند n تا n تا اگر ترتیب قرار گرفتنشان مهم نیست و مهم نباشد. چند طریق می‌توانند در صف قرار بگیرند؟

$\frac{P^n_n}{n!} = 1 = \frac{n!}{n!} \qquad \frac{P^n_r}{r!} = \frac{\frac{n!}{(n-r)!}}{r!} = \frac{n!}{r!(n-r)!}$

$C^n_r = \binom{n}{r}$

مثال ← از ۵ نفر می‌خواهیم یک گروه ۳ تایی انتخاب کنیم؟

مثال ← ۱۷ نفر داریم و می‌خواهیم از بین آن گروه‌ها ۵ تایی و ۷ تایی درست کنیم. داخل گروه‌ها ترتیب مهم نیست اما بین گروه‌ها ترتیب مهم است.

$\frac{17!}{5!\,7!\,5!}$

یا طریق دیگر →

$\binom{17}{5} \times \binom{12}{7} \times \binom{5}{5}$

$\frac{17!}{5!\,12!} \quad \frac{12!}{7!\,5!} \quad \frac{5!}{5!\,0!}$

مثال ← ۱۶ نفر داریم و می‌خواهیم از بین گروه‌های ۵ تایی، ۷ تایی درست کنیم و ترتیب در گروه‌ها مهم نیست.

$\rightarrow \frac{16}{5!\,7!\,1!\,1!\,1!\,1!}$

* جابه‌جایی در ۵ تایی و ۷ تایی مهم نیست

* فرق ترکیب و ترتیب جابه‌جایی است یعنی اهمیت جابه‌جایی.

۰، ۲، ۳، ۴، ۵

مثال ← با اعداد روبرو چند عدد ۳ رقمی با تکرار؟
چند عدد ۳ رقمی بدون تکرار؟

تکرار مجاز هست (تکرار مجاز نیست) = $\underline{۵} \times \underline{۶} \times \underline{۶}$

تکرار مجاز نیست = $\underline{۵} \times \underline{۵} \times \underline{۴} = ۱۰۰$

نکته ← (کد با رقم فرق دارد)

زوج: چند عدد ۳ رقمی می‌توان ساخت که تکرار مجاز نباشد و زوج باشد؟ →

$\underset{۵}{\underline{\quad}} \times \underset{۴}{\underline{\quad}} \times \underset{①}{\underline{۰}}$ حالت

یا

$\underset{۴}{\underline{\quad}} \times \underset{۴}{\underline{\quad}} \times \underset{۱}{\underline{۲}}$

یا

$\underset{۴}{\underline{\quad}} \times \underset{۴}{\underline{\quad}} \times \underset{۱}{\underline{۴}}$

⟹ یا از طریق دیگر: $\underset{۴}{\underline{\quad}} \times \underset{۴}{\underline{\quad}} \times \underset{۲}{\underline{\text{۲ یا ۴}}}$

$۵\times۴ + ۴\times۴\times۲ = ۵۲$

فرد: چند عدد ۳ رقمی می‌توان ساخت که تکرار مجاز نباشد و فرد باشد؟ →

$\underset{۴}{\underline{\quad}} \times \underset{۴}{\underline{\quad}} \times \underset{۳}{\underline{\text{۱ یا ۳ و ۵}}} = ۴۸$

$۵۲ + ۴۸ = ۱۰۰$

چه تعداد از ۳۳۰ بزرگتر است؟ $۳۳۰ < x$

(تکرار مجاز نیست)

$\underset{۲}{\underline{\text{۴ یا ۵}}} \times \underset{۵}{\underline{\quad}} \times \underset{۴}{\underline{\quad}}$

یا

$\underset{۱}{\underline{۳}} \times \underset{۲}{\underline{\text{۴ یا ۵}}} \times \underset{۴}{\underline{\quad}}$

مثال ← ۳ تا پزشک داریم ۴ تا مهندس:

پزشک $P_۳$ مهندس $E_۴$ $\rightarrow (۳+۴)!$ ۷ تایی

۱- به چند طریق می‌توانند در صف قرار بگیرند این ۳ پزشک و ۴ مهندس

۲- به چند طریق می‌توانند کنار هم قرار بگیرند که دکترها کنار هم و مهندس‌ها هم کنار هم (ترتیبشان)

پزشک
↑↑↑
۱ ۲ ۳

مهندس
↓↓↓↓
۱ ۲ ۳ ۴

$۳! \times ۴! \times ۲!$

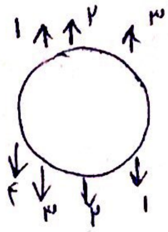
۳- به چند طریق می‌توانند کنار هم قرار بگیرند که یکی در میان پزشک و مهندس باشند

↓ ↑ ↓ ↑ ↓ ↑ ↓
۱ ۱ ۲ ۲ ۳ ۳ ۴

$۴! \times ۳!$

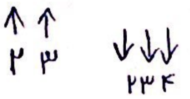
نکته ← اول مهندس است چون تعدادش زیادتر است
اگر تعداد برابر بود فرق و هم نبود

Scanned with CamScanner



۵- به چند طریق می‌توانند دور میز گرد بنشینند

$(7-1)!$



۴- به چند طریق می‌توانند یک دختر و یک پسر خاص کنار هم قرار بگیرند

(خاص ← از قبل انتخاب شده اند)

۶- به چند طریق می‌توانند یک پسر خاص در ۲ دختر خاص

$(5-1)! \times 3!$

۷- از این گروه ۷۰ نفر که می‌خواهیم یک گروه ۴تایی درست و انتخاب کنیم

$\binom{7}{4}$

۸- یک گروه ۴تایی که ۲تا آن حتماً پزشک باشد

$\binom{3}{2} \times \binom{4}{2}$

۹- حداقل ۲تا دکتر باشد یک گروه ۴تایی انتخاب کرده حداقل ۲تا آن دکتر باشد

$\binom{3}{2} \times \binom{3}{2} + \binom{3}{3} \times \binom{4}{1}$

(+ : یا)

(ادامه)

تعاریف: آزمایش تصادفی قبل از انجام آزمایش نتیجه آن مشخص نیست - و اگر در شرایط یکسان انجام شود جواب‌ها متفاوت

$S=\{H, TH, TTH, \ldots\}$

مثال ← پرتاب سکه:
پرتاب تاس
دمای بدن انسان

سکه $S=\{H,T\}$

تاس $S=\{1,2,3,4,5,6\}$

فضای نمونه‌ای را با S نشان می‌دهند. تعداد حالات ممکن یک آزمایش است. مثال ← پرتاب سکه و یا پرتاب تاس یک فضای نمونه‌ای است یا دمای بدن انسان

دمای بدن $S=\{35, 42\}$

مثال ← تعداد کالاهای سالم در یک نمونه ۳تایی از محصولات در یک کارخانه؟

$S=\{NNN, NNN', NN'N, N'NN, NN'N', N'NN', N'N'N, N'N'N'\}$

یا از طریق دیگر $\{0, 1, 2, 3\}$

فضای:
- گسسته
  - متناهی
  - نامتناهی
- پیوسته

احتمال ۳ تا داریم:
- گسسته متناهی
- گسسته نامتناهی
- پیوسته

پیشامد ← هر زیرمجموعه از فضای نمونه را یک پیشامد گویند

○ هرگز اتفاق نمی‌افتد

اتفاق می‌افتد

پیشامد:
- محال
- حتمی

پیشامد مرکب ← پیشامدی که بیش از یک جزء دارد

< یک پیشامد وقتی اتفاق می‌افتد که جواب آزمایش تصادفی جزوی از آن باشد >

مثال ← یک تاس را پرتاب می‌کنیم احتمال آنکه عدد آمده بیش از ۷ باشد؟ ← محال است

احتمال آنکه عدد آمده زوج باشد؟ ← (۲، ۴، ۶) (یک مرکب است)

عدد آمده کمتر از ۲ باشد ← ۱ مجموعه ساده است

عدد آمده کمتر از ۷ باشد ← حتمی است

فقط با یک مطالعه (خوانده)

تعریف مجموعه: جمعیت یا جامعه آماری ← مجموعه یا دسته ای از اشیاء که هر کدام دارای یک خصوصیت مشترک هستند و قسمتی از زیر جمع مجموعه ها اند

انواع مهم مجموعه ها

جبر مجموعه ها:
- ∩ اشتراک
- ∪ اجتماع
- − تفاضل
- Δ تفاضل متقارن

مفهوم متمم و متمم گیری

متمم ها → A, A' ؛ $\varnothing, S$

قاعده دمورگان

$(A\cap B)' = A' \cup B'$

$(A \cup B)' = A' \cap B'$

عضویت و زیر مجموعه بودن ← $\in \quad \subseteq$

ناسازگاری و ناسازگار بودن ←

$A \cap B = \varnothing \iff$ A و B ناسازگارند: اشتراکی ندارند

تعداد عضوهای یک مجموعه ←

$n(A \overset{\text{اجتماع}}{\cup} B) = n(A) + n$

$n(A-B) = n(A) - n(A\cap B)$

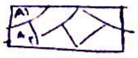
$n(S) = n(A) + n(A)'$

$n(A \Delta B) = n(A\cup B) - n(A \cap B)$ ← (تفاضل) تفاضل متقارن

یا $= n(A-B) + n(B-A)$

افراز ها ←

$\cdot S$



شرط اول $= A_i \cap A_j$ ، $i \neq j$

شرط دوم $= A_1 \cup A_2 \cup \cdots \cup A = S$

Scanned with CamScanner

احتمال ← شخصی / هم‌شانسی / * بر اساس فراوانی نسبی

احتمال همراه

احتمال بر اساس فراوانی نسبی

$\frac{f_n(A)}{n} = r_n(A)$ فراوانی خصوصیات A در یک نمونه n تایی است

اگر تقسیم بر n کنیم می‌شود فراوانی نسبی $r_n(A)$

$\lim_{n\to\infty} r_n(A) = P(A)$ اگر آزمایش هر چقدر بیشتر باشد وقتی $n \to \infty$ (بی‌نهایت) اگر به سمت عدد ثابتی برود P(A) احتمال

احتمال بر اساس تابع ← ۳ شرط اساسی دارد

$P: S \longrightarrow R^{[0,1]}$

اصول موضوع احتمال

① $A \subseteq S$ ، $0 \leq P(A) \leq 1$

② $P(S) = 1$ (حتمی) و $P(\emptyset) = 0$ (محال)

③ مهم: $P(\bigcup_{i=1}^{k} A_i) = \sum_{i=1}^{k} P(A_i)$ (احتمال اجتماع)

$A_i \cap A_j = \emptyset$

مهم: فضاها مدل احتمالی روی فضای نمونه‌ای گسسته و متناهی

فرض کنیم S یک مجموعه متناهی و غیر تهی باشد و اعضای آن e عضوی از فضای نمونه باشد

طوری که $P(e_i) = P_i$ ، $0 \leq P_i \leq 1$ و مجموع آن $\sum_{i=1}^{n} P_i = 1$ مدل بدست آمده را به آن فضا مدل

احتمال روی نمونه‌ای گسسته و متناهی است

$S = \{e_1, e_2, ---- e_n\}$

$P(e_i) = P_i \quad 0 \leq P \leq 1$

$\sum_{i=1}^{n} P = 1$

حالت	e_1	e_2	----	e_n
احتمال	P_1	P_2		P_n

$\sum^{n} p_i = 1$ ، $0 \leq P_i$

احتمال وقوع پیشامد A چقدر است ← مثلا $A = \{e_1, e_8\}$

$P(A) = P(\{e_1, e_8\})$

$= P(\{e_1\} \cup \{e_8\})$

$= P(\{e_1\}) + P(\{e_8\})$

$= P_1 + P_8$

هم‌شانس است: $P_i = P$ ، $0 \leq P$ ، $\sum_{i=1}^{k} P = 1 \longrightarrow nP = 1$ ، $P = \frac{1}{n}$

$P(A) = P_1 + P_8 = 2P = \frac{2}{8} = \frac{n(A)}{n(S)} \longleftarrow$

Scanned with CamScanner

→ در هم‌شانسی باید برابر و برابر P هستند

(نکته) فرمول‌های هم‌شانسی برای بدست آوردن احتمال: تعداد حالات مساعد به تعداد کل حالات ممکنه

مثال ← یک تاس سالم را پرت می‌کنیم و مدل احتمال آن را بدست و احتمال زوج بودن آن را حساب کنید

- اگر این تاس دستکاری شده باشد به طوری که احتمال زوج بودن ۲ برابر فرد بودن باشد

(گسسته و متناهی است)

$S = \{1, 2, 3, 4, 5, 6\}$

حالت‌ها	1	2	3	4	5	6	
احتمال	p	p	p	p	p	p	$6p=1$, $p=\frac{1}{6}$

احتمال زوج بودن $P(\text{زوج}) = P(\{2,4,6\}) = \frac{3}{6}$

قسمت دوم سوال

$P(\text{زوج}) = 2 \qquad P = \text{فرد}$

احتمال فرد بودن $1 = P$

احتمال زوج بودن $= 2P$

شرط داریم

1	(2)	3	(4)	5	(6)	
p	2p	p	2p	p	2p	$9p=1$
1/9	2/9	1/9	2/9	1/9	2/9	$p=\frac{1}{9}$

$$P(\{2,4,6\}) = P_2 + P_4 + P_6 = \frac{2}{9} + \frac{2}{9} + \frac{2}{9} = \frac{6}{9}$$

مهم ← احتمال: اگر احتمال هر عدد متناسب با همان عدد باشد

حالت	1	2	3	4	5	6	
احتمال	1p	2p	3p	4p	5p	6p	$21p=1$
	1/21	2/21	3/21	4/21	5/21	6/21	$p=\frac{1}{21}$

احتمال زوج بودن چقدر است

$$P(\text{زوج}) = P(\{2,4,6\}) = \frac{2}{21} + \frac{4}{21} + \frac{6}{21}$$

Scanned with CamScanner

مثال مهم در مسابقه ورزشی

مثال ← احتمال آنکه افراد ۷ نفره که در صف قرار گرفته‌اند بیا در میان دکتر و مهندس باشند (مدل هم شناسی است)

$n(S) = ۷!$

$n(A) = ۴! \times ۳!$

$p(A) = \frac{۴! \times ۳!}{۷!}$

احتمال آنکه ۲ نفر خاص دور میز خاص کنار هم باشند

$n(A) = ۵! \times ۲!$

$n(S) = ۶!$

احتمال آنکه در یک گروه ۴ تایی حداقل ۲ تا دکتر باشند

$$\frac{\binom{۳}{۲}\times\binom{۴}{۲} + \binom{۳}{۳}\times\binom{۴}{۱}}{\binom{۷}{۴}}$$

سوال ← پمپ بنزین داریم تحت بازرسی هستند و ۲ تا عیب دارند و داریم تحت بازرسی قرار داریم

	PK نشت	Q کار نکردن دستگاه	KQ نشت و کار نکردن دستگاه
A	۰/۱	۰/۸	۰
B	۰/۱	۰/۱	۰/۰۸

احتمال معطل شدن کدام پمپ بنزین در بازرسی بیشتر است (حداقل یک عیب را برحسب کنم) (بیراحتمال)

$P(\text{حداقل عیب}) = P(K \cup Q) = P(K) + P(Q) - P(K \cap Q)$ حداقل عیب

آنالیز

دومی $۰/۱ + ۰/۸ - ۰ = ۰/۹$

اولی $۰/۱ + ۰/۱ - ۰/۰۸ = ۰/۱۲$ ✓

<احتمال هیچ وقت متوقف نشود>

۱ احتمال آنکه اولی به احتمال یک عیب معطل شود ← $۰/۱ + ۰/۰۸ = ۰/۱۸$

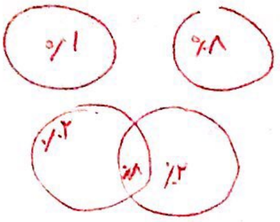
۲ احتمال آنکه دومی به احتمال یک دلیل معطل شود چقدر است ← تفاضل متقارن

$P(A \Delta B) = P(A-B) + P(B-A) = P(A \cup B) - P(A \cap B)$

اولی $P(K \Delta Q) = P(K \cup Q) - P(K \cap Q) = ۰/۱۸ - ۰ = ۰/۱۸$

دومی $P(K \Delta Q) = ۰/۱۲ - ۰/۰۸ = ۰/۰۴$

Scanned with CamScanner



اولی $P(K\Delta Q) = P(K \cup Q) - P(K \cap Q)$ = ۰/۱۸ − ۰ = ۰/۱۸

دومی $P(K \Delta Q)$ = ۰/۱۲ − ۰/۰۸ = ۰/۰۴

← احتمال آن که پمپ بنزین دومی به خاطر نشت تعطیل شود ←

$P(K - Q) = P_1(K) - P(K \cap Q)$

= ۰/۱ − ۰/۰۸ = ۰/۰۲

← احتمال آنکه پمپ بنزین دوم به خاطر قطع شدن تعطیل شود

$P(Q - K)$ = ۰/۱ − ۰/۰۸ = ۰/۰۲

مثال ←

(گسسته نامتناهی) حذف شده نیست

* (گسسته نامتناهی) پرتاب یک سکه و اینقدر پرتاب می‌کنیم تا شیر بیاید

با یک تاس را اینقدر پرتاب می‌کنیم تا اولین بار ۶ بیاید مدل احتمال را بنویسید

$S = \{e_1, e_2, \ldots\}$

$P(e_i) = p_i$

پیامدها $e_1 \quad e_2 \quad \ldots$

مثال ←

تاس را آنقدر پرتاب کنیم تا اینکه اولین بار ۶ بیاید موارد احتمال آن را بنویسید؟

تعداد دفعات فرد و زوج برابر است یا نه ← کدام بیشتر؟

موارد احتمال $S = \{1, 2, 3, 4, \dots\}$

	۱	۲	۳	۴	
	$\frac{1}{6}$	$\frac{5}{6}\times\frac{1}{6}$	$\frac{5}{6}\times\frac{5}{6}\times\frac{1}{6}$	$(\frac{5}{6})^3\times\frac{1}{6}$	$\dots$

آیا جمع آنها یک می‌شود؟ باید یک بشود؟

زوج

$$\left(\frac{1}{6}+\frac{1}{5}\cdot\frac{1}{6}+\left(\frac{5}{6}\right)^2\times\frac{1}{6}+\dots\right)=\frac{1}{6}\left(1+\frac{5}{6}+\left(\frac{5}{6}\right)^2+\dots\right)$$

فاکتورگیری

حد مجموع

$$=\frac{1}{6}\left(1\times\frac{1-(\frac{5}{6})^\infty}{1-(\frac{5}{6})}\right)=1$$

فرد

$$\frac{1}{6}+\left(\frac{5}{6}\right)^2\times\frac{1}{6}+\left(\frac{5}{6}\right)^4\times\frac{1}{6}+\dots$$

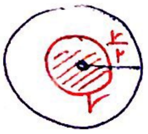
$$\frac{1}{6}\left(1+\left(\frac{5}{6}\right)^2+\left(\frac{5}{6}\right)^4+\dots\right)$$

$$\frac{1}{6}\left(1\ \frac{1-((\frac{5}{6})^2)^\infty}{1-(\frac{5}{6})^2}\right)=\frac{\frac{1}{6}}{(\frac{5}{6})^2}$$

فرمول $\frac{1-q^\infty}{1-q}$

** پیوسته ها > (در داده های پیوسته برابر است با صفر) : احتمال وقوع نقطه صفر است . ← احتمال وقوع نقطه ← نکته
(این بازه ها هستند که احتمال وقوع پیدا می کنند)

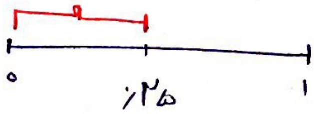
$$P(A) = \frac{\text{طول مساعد}}{\text{طول کل}} = \frac{\text{حجم مساعد}}{\text{حجم کل}} = \frac{\text{دمای مساعد}}{\text{دمای کل}} = \frac{\text{وزن مساعد}}{\text{وزن کل}}$$



مثال ← یک دایره به شعاع r و یک نقطه درون دایره هست
احتمال آن که این نقطه به مرکز نزدیک تر از محیط دایره باشد چقدر است

$$P(A) = \frac{\pi (\frac{r}{2})^2}{\pi r^2} = \frac{1}{4}$$

مثال ← یک عدد بین صفر و یک انتخاب احتمال آن که عدد کمتر از ۲۵٪ باشد؟



جواب ←
$$\frac{0/25 - 0}{1 - 0} = \frac{1}{4}$$

یا بین ۲۵٪ و ۷۵٪ باشد!؟ —

مثال ← خانواده‌ای دو تا فرزند دارد حالات ممکن که این فرزندها می‌توانند داشته باشند چه قدر است و تکرار مجاز است

$0,0$

$2 \times 2 = 4$

$S = \{GG, GB, BG, BB\}$

اگر این خانواده حداقل یک فرزند دختر داشته باشد چه فضای نمونه وجود دارد؟

$B = \{GG, GB, BG\}$

* اگر فضای نمونه‌ای تغییر کند احتمالات هم تغییر می‌کند

تعریف احتمال شرطی (احتمال رخ دادن پیشامد (A) به شرط آن می‌باشد که پیشامد B رخ داده باشد را احتمال شرطی گویند.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \qquad P(B) \neq 0$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

اگر جای A و B را عوض کنیم

$$P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

ضرب احتمالات

مثال ← ما یک خانواده‌ای دو فرزند داریم و می‌دانیم که این خانواده حداقل یک فرزند دختر دارد احتمال اینکه این هر دو فرزند دختر باشند چه قدر است؟

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}$$

مثال ←

یک تاس را پرتاب کردیم می‌دانیم عددی که آمده بیشتر از ۳ است احتمال اینکه زوج باشد چه قدر است

$S = \{1, 2, 3, 4, 5, 6\}$

$B = \{4, 5, 6\}$

$A = \{4, 6, 2\}$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\{4, 6\})}{P(B)} = \frac{\frac{2}{6}}{\frac{3}{6}} = \frac{2}{3}$$

مثال ← فرض کنید ۳ تا پیشامد A و B و C اتفاق افتاده است به طوری که اول A و دوم B و سوم C اتفاق افتاده است و ضرب احتمالات آن؟

$$P(A \cap B \cap C) = P(A) \cdot P(B|A) \cdot P(C|A, B)$$

مثال ← $3W$ $4B$ جعبه‌ای است که ۳ تا مهره سفید و ۴ تا مهره سیاه دارد ۲ مهره با جایگذاری انتخاب کنیم احتمال اینکه هر دو مهره سفید باشد چه قدر است؟

با جایگذاری (با تکرار): $P(w_1, w_2) = \frac{3}{7} \times \frac{3}{7}$

بدون جایگذاری: $P(w_1, w_2) = P(w_1) \cdot P(w_2|w_1) = \frac{3}{7} \times \frac{2}{6}$

یا → $\dfrac{\binom{3}{2}\binom{4}{0}}{\binom{7}{2}}$

Scanned with CamScanner

۲۰ — قانون ضرب احتمالات

← اگر اولی سفید و دومی سیاه و سومی سفید ؟

$$P(\omega_1, B_2, \omega_3) = P(\omega_1) \cdot P(B_2|\omega_1) \cdot P(\omega_3|B_2, \omega_1) = \frac{3}{7} \times \frac{4}{6} \times \frac{2}{5}$$

— با جایگذاری شرط احتمال نمی‌شود.
— بدون جایگذاری شرط احتمال می‌شود.

مثال ← من یک جعبه دارم مهره اول در می‌آورم بیرون می‌گذارم و دومی را در می‌آورم احتمال اینکه مهره دوم سیاه باشد چقدر است؟ (بدون جایگذاری)

$$P(B_2) \rightarrow P(B_1, B_2) + P(\omega_1, B_2)$$

$$= P(B_1) \cdot P(B_2|B_1) + P(\omega_1) \cdot P(B_2|\omega_1)$$

$$= \frac{4}{7} \times \frac{3}{6} + \frac{3}{7} \times \frac{4}{6}$$

$$= \frac{4}{7}\left(\frac{3}{6} + \frac{3}{6}\right) = \frac{4}{7}$$

نکته مهم ← مفهوم استقلال

دو پیشامد A و B مستقل اند اگر وقوع یکی بر احتمال دیگری اثر نداشته باشد مستقل است و اگر اثر داشته باشد از هم مستقل نیست

$$P(A \cap B) = P(A) \cdot P(B)$$

$$P(A|B) = P(A)$$

$$P(B|A) = P(B)$$

*

مفهوم ناسازگاری ←
1- $A \cap B = \varnothing$ ناسازگارند
2- $P(A \cap B) = 0$

← ۲ شرط مهم است

اگر A و B از هم مستقل اند A و B → A, B′
A′ و B
A′, B′
سازگارند

مفهوم سازگاری ← A و B مستقل

$$P(A \cap B) = P(A) \cdot P(B)$$

$$P(A|B) = P(A)$$

$$P(B|A) = P(B)$$

نکته ←
— اگر پیشامدها ۲ به ۲ مستقل باشند و ۳ تایی و ۴ تایی و بیشتر از هم مستقل اند — و در ناسازگاری ۲ تایی و ۳ تایی و ... با هم ناسازگارند
اما این خاصیت در استقلال برقرار نیست

$$P(A \cap B) = P(A) \cdot P(B)$$

$$P(A \cap C) = P(A) \cdot P(C)$$

$$P(C \cap B) = P(C) \cdot P(B)$$

$$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$$

Scanned with CamScanner

مثال ← ۲ تاس را می‌اندازیم و پیشامدهای A و B و C را تعریف می‌کنیم نشان دهید که ۲ به ۲ مستقل اند و ۳ تایی هم مستقل نیستند؟

A = اولین تاس ۳ بیاید

B = دومین تاس ۴ بیاید

C = مجموع دو تاس ۷ بشود

$A = \{(3,1),(3,2),(3,3),(3,4),\dots\}$ $\quad n(A) = 6$

$B = \{(1,4),(2,4),(3,4),\dots\}$ $\quad n(B) = 6$

$C = \{(1,6),(2,5),(3,4),\dots\}$ $\quad n(C) = 6$

حالت ۲ تایی

شرط استقلال

$P(A\cap B) = P(A).P(B) \rightarrow \frac{1}{36} = \frac{6}{36}\cdot\frac{6}{36}$ (کل حالات)

$P(A|B) = P(A) \rightarrow \frac{1}{6} = \frac{6}{36}$ (حالات مطلوب / حالات کل)

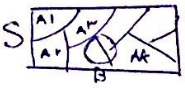
حالت ۳ تایی

$P(A\cap B\cap C) = P(A).P(B).P(C)$

$\frac{1}{36} \neq \frac{6}{36}\cdot\frac{6}{36}\cdot\frac{6}{36}$ (مستقل نیستند) — مستقل نیستند.

$P(C|A,B) = P(C)$

$1 = \frac{6}{36}$ (اشتراک دارند)



(احتمال کل - بیز) >

اگر یک فضای نمونه‌ای بدون اشتراک توسط A ها افراز می‌شود (شرط افراز) ←

شرط افراز:

1- $A_i \cap A_j = \emptyset \quad i \neq j$

2- $\bigcup_{i=1}^{K} A_i = 1$ (اجتماع)

مثلاً اگر یک پیشامد B داشته باشیم آیا افرازهای S آیا B را افراز می‌کنند؟ نه (مثل اینکه نقاشی کنیم و با رنگ دو بار کنار هم قرار دهیم) پس هر زیر پیشامدی توسط افرازهای S افراز می‌شود.

$B \subseteq S$

$P = (B)$

$B = \bigcup_{i=1}^{k} (B\cap A_i)$ (اجتماع)

$P(B) = P(\bigcup_{i=1}^{k} (B\cap A_i))$

قضیه احتمال کل ←

$= \sum_{i=1}^{k} P(B\cap A_i)$

$= \sum_{i=1}^{k} P(A_i).P(B|A_i)$

Scanned with CamScanner

قضیه بیز ← ما بالا به دنبال احتمال مثلاً A_3 یا A_4 هستیم به شرطی B اتفاق افتاده باشد و دنبال جزئی از کل هستیم

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i)\cdot P(B|A_i)}{\sum_{i=1}^{k} P(A_i)\cdot P(B|A_i)}$$

جزء به کل

مثال ← ۳ تا کارخونه داریم که یک محصول خاص را تولید می‌کنند و شرکتی است که از این ۳ تا کارخانه خرید می‌کند و این کارخانه‌ها اعلام کردند که چند درصد از کالاهاشون معیوب هستند و درصد خوب از ۳ تا کارخانه خرید می‌کنند شرکت در کارخونه‌ای به هر دلیلی تعطیل شود عمل می‌کنه از کارخونه‌های دیگه درصدی که لازم است خرید کنم

	$P(N'\|A_i)$ N'	$P(A_i)$ میزان خرید
A	%۵	%۸۰
B	%۱۵	%۱۵
C	%۲۰	%۵

۱- احتمال اینکه یک کالا معیوب باشد چقدر است (اینکه کالا را انتخاب و احتمال اینکه ...) کل احتمال

۲- اگر این کالا معیوب باشد احتمال اینکه از کارخونه سوم باشد چقدر است؟

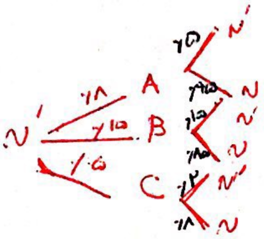
قسمت اول سوال جواب →

$$P(N') = P(N',A) + P(N',B) + P(N',C)$$

$$= P(A)\cdot P(N'|A) + P(B)\cdot P(N'|B) + P(C)\cdot P(N'|C)$$

$$= 0/8 \times 0/05 + 0/15 \times 0/15 + 0/05 \times 0/2$$

رسم نمودار شاخه درختی



قسمت دوم سوال (بیز) →

$$P(C|N') = \frac{P(C)\cdot P(N'|C)}{0/8\times 0/05 + 0/15 \times 0/15 + 0/05 \times 0/2}$$

Scanned with CamScanner

مثال: احتمال اتمی $P(X=1{,}5)$ چقدر است؟

احتمال هر نقطه صفر است $P(X=1{,}5)=0$

مثال ← احتمال اتمی $P(x=1)$ چقدر است؟

$$P(x=1)=\frac{3}{8}$$

$$=F_X(1^+)-F_X(1^-)$$

(حد راست) (حد چپ)

$$=\frac{4}{8}-\frac{1}{8}=\frac{3}{8}$$

مثال ← احتمال را بدست آورید $P(1\leq x<3)$

$$P(1\leq x<3)=P(x=1)+P(x=2)=\frac{6}{8}$$

$$F(3^-)-F(1^-)=\frac{7}{8}-\frac{1}{8}=\frac{6}{8}$$

مثال ← احتمال را بدست آورید $P(0<X\leq 2)$

$$P(0<X\leq 2)=P(x=1)+P(X=2)=\frac{6}{8}$$

تابع توزیع $F(2^+)-F_X(0^+)=\frac{7}{8}-\frac{1}{8}=\frac{6}{8}$

مثال ← احتمال را بدست آورید $P(0\leq x<2)$

$$P(0\leq x<2)\rightarrow\frac{4}{8}$$

احتمال: $P(x=0)+P(x=1)=\frac{1}{8}+\frac{3}{8}=\frac{4}{8}$

تابع توزیع: $F(2^-)-F(0^-)=\frac{4}{8}-0=\frac{4}{8}$

مثال ← احتمال را بدست آورید؟ $P(x>1)$

$$P(x>1)\rightarrow P(x=2)+P(x=3)=\frac{4}{8}$$

تابع توزیع نیست و باید تبدیل کنیم ← متمم

تابع توزیع $=1-P(x\leq 1)$

$$=1-F_X(1^+)=1-\frac{4}{8}=\frac{4}{8}$$

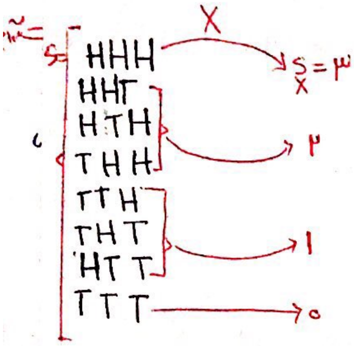
Scanned with CamScanner

(متغیر تصادفی - تابع چگالی - تابع توزیع - جدول توزیع احتمال)

متغیر تصادفی ← یک تابع است از $X: S \longrightarrow \mathbb{R}$ (مجموع اعداد حقیقی) (اگر گسسته باشد)

★ مثال ۱ ← X = تعداد شیرها در ۳ بار پرتاب سکه (گسسته است)

$S = \{0, 1, 2, 3\}$



$S = \{0, 1, 2, 3\}$

★ مثال ۲ ← دو تا تاس را پرتاب می‌کنیم

X = مجموع اعداد ۲ رو در دو تاس

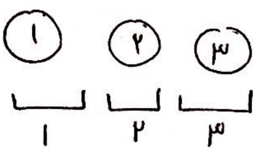
$S_A = \{2, 3, 4, 5, ..., 12\}$

X = مجموع اعداد ۲ تاس

$S_X = \{2, 3, 4, 5, ..., 12\}$

$S = \{(1,1),(1,2) ... (1,6);\ (2,1), ..., (2,6);\ ...;\ (6,1), ..., (6,6)\}$

★ مثال ۳ ← ۳ تا جای خالی داریم و ۳ تا مهره و در هر جای خالی فقط یک مهره قرار می‌گیرد و مهره ها شماره دارند و جاها هم شماره دارند



X = تعداد جورها (یعنی شماره جعبه و مهره یکی باشد)

تعداد حالت ها $S = \{123, 132, 213, 231, 312, 321\}$

(۳ جور دارد، ۱، ۱، ۰، ۰، ۱) — ۶ حالت می‌تواند داشته باشد

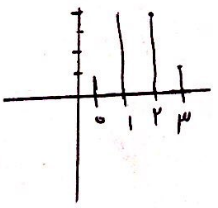
$S_X = \{0, 1, 3\}$

— جدول توزیع احتمال - تابع چگالی - تابع توزیع را حالات خواسته شده ... را حساب کنید برای مثال ۱ بالا $P(X=x)$

X	0	1	2	3	
$P(X=x)$	1/8	3/8	3/8	1/8	1

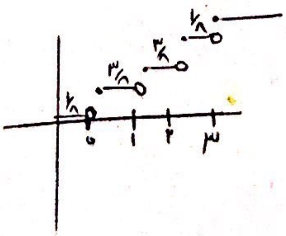
جدول توزیع احتمال

تابع چگالی
$$f_X(x) = P(X=x) = \begin{cases} 1/8 & x=0 \\ 3/8 & x=1 \\ 3/8 & x=2 \\ 1/8 & x=3 \\ 0 & \text{سایر نقاط} \end{cases}$$



تابع توزیع
$$F_X(x) = \begin{cases} 0 & x<0 \\ 1/8 & 0 \le x < 1 \\ 4/8 & 1 \le x < 2 \\ 7/8 & 2 \le x < 3 \\ 1 & x \ge 3 \end{cases}$$

$P(X \le x)$



Scanned with CamScanner

نکته ← در داده‌های پیوسته اول باید تابع توزیع را بدست آوریم

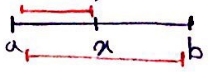
نکته ← رابطه بین تابع توزیع و چگالی مشتق و انتگرال است
و از تابع توزیع مشتق بگیریم تابع چگالی بدست می‌آید

← گسسته‌ها قابل شمارش بودند

پیوسته‌ها ← (وزن) (زمان)

مثال ← متغیر تصادفی X:

$X : a, b$ انتخاب یک x در بازه a و b به طوری که $a<b$



تابع چگالی $P(X=x)=0$

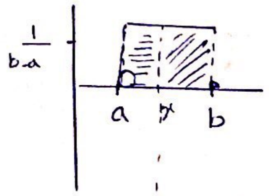
در داده‌های پیوسته احتمال وقوع هر نقطه‌ای صفر است و همیشه صفر است

تابع توزیع → $F_X(x) = P(X \leq x) = \begin{cases} 0 & x<a \\ \dfrac{x-a}{b-a} & a \leq x < b \\ 1 & x \geq b \end{cases}$

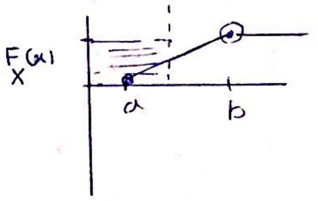
طول ... / تناسب مشخص است

$\frac{1}{b-a}(x-a)$ → مشتق $f_X(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{سایر} \end{cases}$ (تابع چگالی بدست آمد)

تابع چگالی

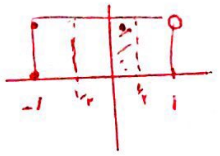


تابع توزیع



مثال ← $a=-1 \quad b=1$

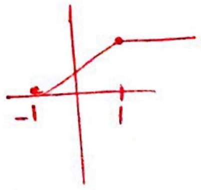
تابع چگالی $f_X(x) = \begin{cases} \frac{1}{1-(-1)} & -1 \leq x < 1 \\ 0 & \end{cases}$ 0.5



احتمال آنکه $x=\frac{1}{2}$ ؟ برابر است با صفر

نکته + احتمال وقوع نقطه هر داده برابر صفر است در داده پیوسته.

$F_X(x) = \begin{cases} 0 & x<-1 \\ \dfrac{x-(-1)}{1-(-1)} & -1 \leq x < 1 \\ 0 & x \geq 1 \end{cases}$



مثال ← $P(x=\frac{1}{2})=0$

تابع توزیع $F_X(\frac{1}{2}) - F_X(0)$

$\dfrac{\frac{1}{2}-(-1)}{1-(-1)} - \dfrac{0-(-1)}{1-(-1)}$

$\dfrac{1/5}{2} - \dfrac{1}{2} = \dfrac{0/5}{2} = \dfrac{1}{4}$

تابع چگالی $P(x=\frac{1}{2})=0$

$P(0 \leq x \leq \frac{1}{2})$

$= \int_0^{\frac{1}{2}} f(x)\,dx$

$= \int_0^{\frac{1}{2}} \frac{1}{2}dx = \frac{1}{2}x \Big|_0^{\frac{1}{2}}$

$= \frac{1}{2}(\frac{1}{2}-0) = \frac{1}{4}$

Scanned with CamScanner

مثال ← $P(x=\frac{1}{2})=0$

$P(\frac{1}{2}<x<-\frac{1}{2}) \to$

تابع توزیع

$F_X(\frac{1}{2})-F_X(-\frac{1}{2})$

$\frac{\frac{1}{2}-(-1)}{} - \frac{\frac{1}{2}-(-1)}{1-(-1)}$

$= \frac{1/5}{2} - \frac{0/5}{2} = \frac{1}{2}$

مثال ← $P(X=\frac{1}{2})=0$

$P(x>0/75)$

با دو استفاده انتگرال

$= \int_{0/75}^{\alpha} f_X(x)\,dx = \int_{/75}^{1} \frac{1}{2}\,dx + \int_{1}^{\alpha} = dx$ راه اشتباه است

تابع چگالی

$= \int_{/75}^{1} \frac{1}{2}\,dx = \frac{1}{2}(x\Big|_{/75}^{1})$

$= \frac{1}{2}(1-0/75)$

تابع توزیع

$F_X(1)-F_X(0/75) = \frac{1-0/75-(-1)}{2}$

$= 1-\frac{1/75}{2}$

Scanned with CamScanner

سوالات

مثال ص ۳۱۱

$$f_X(x)=\begin{cases} kx & 0<x<2 \\ 0 & \text{o.w} \end{cases}$$

۱- k را طوری بدست آورید $k=?$

۲- تابع توزیع را بدست آورید $F_X(x)=?$

۳- احتمالات زیر را حساب کنید $P(x>1)$

۴- $P(1.5<x<2.5)$

با دو شرط

هر دو باید معتبر

1) $f_X(x)>0$

2) → $\int_{-\infty}^{+\infty} f_X(x)\,dx=1$

$kx>0$

$0<x$ $\quad k>0$

$$\int_{-\infty}^{+\infty} f_X(x)\,dx=1$$

$$1=\int_0^2 kx\,dx=k\left(\frac{x^2}{2}\Big|_0^2\right)=k\left(\frac{2^2}{2}-0\right)=2k=1\longrightarrow k=\frac{1}{2}$$

تابع توزیع قسمت دوم سوال

$$\to F_X(x)=\int_{-\infty}^{x} f_X(x)\,dx=\int_0^x \frac{x}{2}\,dx=\frac{1}{2}\int_0^x x\,dx$$

$$\frac{1}{2}\left(\frac{x^2}{2}\Big|_0^x\right)=\frac{1}{2}\left(\frac{x^2}{2}-0\right)$$

$$=\frac{x^2}{4}$$

تابع توزیع

$$\to F_X(x)=\begin{cases} 0 & x<0 \\ \frac{x^2}{4} & 0\leq x<2 \\ 1 & x\geq 2 \end{cases}$$

$$P(x>1)=\int_1^{\infty} f_X(x)\,dx=\int_1^2 \frac{x}{2}\,dx+\int_2^{\infty} 0=\frac{1}{2}\left(\frac{x^2}{2}\Big|_1^2\right)=\frac{1}{2}\left(\frac{2^2}{2}-\frac{1}{2}\right)$$ قسمت سوم مسئله

تا از طریق دیگر $P=(x>1)\to =1-P(x\leq 1)=1-F_X(1)=1-\frac{(1)^2}{4}$

قسمت چهارم مسئله

$$P(1.5<x<2.5)\to\int_{1.5}^{2.5} f_X(x)\,dx=\int_{1.5}^{2}\frac{1}{2}x\,dx+\int_2^{2.5} 0\,dx=\frac{1}{2}\int_{1.5}^{2}x\,dx=\frac{1}{2}\left(\frac{x^2}{2}\Big|_{1.5}^{2}\right)=$$

$$=\frac{1}{2}\left(\frac{2^2}{2}-\frac{1.5^2}{2}\right)$$

از طریق تابع توزیع $=P(1.5<x<2.5)\to F_X(2.5)-F_X(1.5)=1-\frac{1.5^2}{4}$

Scanned with CamScanner

مثال ← $a=1 \quad b=3$

تابع چگالی → $f_X(x) = \begin{cases} \frac{1}{3-1} & 1 \leq x \leq 3 \\ 0 & \text{o.w} \end{cases}$

$$F_X(x) = \begin{cases} 0 & x < 1 \\ \frac{x-1}{3-1} & 1 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$

مثال ← $P(x \geq 1) = 1$

$$P(1{,}5 \leq x < 2{,}5) = \int_{1{,}5}^{2{,}5} \frac{1}{2} \cdot dx = \frac{1}{2} x \Big|_{1{,}5}^{2{,}5} = \frac{2{,}5 - 1{,}5}{2} = \frac{1}{2}$$

$$P(X \geq 1{,}5) = \int_{1{,}5}^{\infty} f_X(x)\,dx$$

$$= \int_{1{,}5}^{\infty} \frac{1}{2} dx = \int_{1{,}5}^{3} \frac{1}{2} dx + \int_{3}^{\infty} f_X(x)\,dx$$

$$= \frac{1}{2} \; x \Big|_{1{,}5}^{3} = \frac{3-1{,}5}{2} = \frac{1{,}5}{2}$$

مثال ← $P(1{,}1 < x < 2{,}3)$

$$= \int_{1{,}1}^{2{,}3} \frac{1}{2} dx = \frac{2{,}3 - 1{,}1}{2}$$

$$F_X(2{,}3) - F_X(1{,}1)$$

$$= \frac{2{,}3 - 1}{2} - \frac{1{,}1 - 1}{2}$$

Scanned with CamScanner

سوال امتحانی

مثال ← تابع چگالی زیر را در نظر بگیرید مقدار k را طوری به دست آورید که این تابع چگالی شود

تابع توزیع آن را به دست آورید

احتمال $P(\frac{1}{2}<x<\frac{5}{4})$

با داشتن این ۲ شرط ارائه شده →

۱) $f_X(x) > 0 \rightarrow kx^2 > 0,\ k > 0$

۲) $\int f_X(x)\,dx = 1$

۳) $kx^2 > 0$, $k > 0$

جواب → $\int_0^1 kx^2\,dx = 1$ (شرط اول)

$$k\left(\frac{x^3}{3}\Big|_0^1\right) = 1$$

$$k\left(\frac{1-0}{3}\right) = 1$$

$$\frac{k}{3} = 1 \qquad k = 3$$

تابع توزیع

$$F_X(x) = \int_{-\infty}^{x} f_X(x)\,dx = \int_0^x 3x^2\,dx$$

$$\left(3\frac{x^3}{3}\Big|_0^x\right) = x^3$$

$$F(x) = \begin{cases} 0 & x<0 \\ x^3 & 0\leq x<1 \\ 1 & x>1 \end{cases}$$

قسمت سوم مسئله

$$P(\tfrac{1}{2}<x<\tfrac{5}{4}) \rightarrow P(\tfrac{1}{2}<x<1) + P(1<x<\tfrac{5}{4})$$

($P(1<x<\frac{5}{4}) = 0$)

$$= \int_{1/2}^{1} 3x^2\,dx = 3\frac{x^3}{3}\Big|_{1/2}^{1} = (1)^3 - (\tfrac{1}{2})^3$$

یا از این طریق حل

$$F_X(1) - F_X(\tfrac{1}{2})$$

$$= (1)^3 - (\tfrac{1}{2})^3$$

Scanned with CamScanner

$f_X(x) = \begin{cases} \dots \\ 0 \end{cases}$ ... 0.5

(مثال)
1- $k = ?$
2- $F_X(x)$
3- $P(x > \frac{1}{2})$
$P(x < 1.5)$
$P(0.75 < x < \frac{5}{4})$

شرط (1) $\rightarrow k(x+1) \geq 0$

$0 < x < 2$

$1 < x+1 < 3$

$k > 0$

شرط (2) $\rightarrow \int_0^2 k(x+1) = k\int_0^2 (x+1) = k\left(\frac{x^2}{2} + x\right)\Big|_0^2 =$

$= k\left(\frac{4}{2} + 2\right) - 0 = 4k = 1 \quad k = \frac{1}{4}$

نکته: اگر از تابع توزیع مشتق بگیریم می‌شود تابع چگالی و اگر از تابع چگالی انتگرال بگیریم می‌شود تابع توزیع

تابع چگالی $\quad F_X(x) = \int_{-\infty}^{x} f_X(x)\,dx = \int_0^x \frac{1}{4}(x+1)\,dx = \frac{1}{4}\left(\frac{x^2}{2} + x\right)\Big|_0^x = \frac{1}{4}\left(\frac{x^2}{2} + x\right)$

تابع توزیع $\quad F_X(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{4}\left(\frac{x^2}{2} + x\right) & 0 \leq x < 2 \\ 1 & x \geq 2 \end{cases}$

قسمت سوم مسئله $\quad P(x > \frac{1}{2}) \longrightarrow \int_{\frac{1}{2}}^{2} \frac{1}{4}(x+1)\,dx + 0$

$= \frac{1}{4}\left|\left(\frac{x^2}{2} + x\right)\right|_{\frac{1}{2}}^{2}$ ....

$P(x < 1.5) \longrightarrow F_X(1.5) = \frac{1}{4}\left(\frac{(1.5)^2}{2} + 1.5\right)$

$P(0.75 < x < \frac{5}{4}) \longrightarrow = \int_{0.75}^{\frac{5}{4}} f_X(x)\,dx = \frac{1}{4}\left(\frac{(\frac{5}{4})^2}{2} + \frac{5}{4}\right) - \frac{1}{4}\left(\frac{(0.75)^2}{2} + 0.75\right)$

$P(x > \frac{1}{2}) = 1 - P(x \leq \frac{1}{2}) = 1 - F(\frac{1}{2}) = 1 - \frac{1}{4}\left(\frac{(\frac{1}{2})^2}{2} + \frac{1}{2}\right)$

Scanned with CamScanner

‹متغیر تصادفی توام ۲ متغیره›

شروط

① $P(X=x, Y=y) \geq 0$

② $\sum_y \sum_x P(X=x, Y=y)=1$

③ $P(A \cap B)=P(A) \cdot P(B)$

$P(X=x, Y=y)=P(X=x) \cdot P(Y=y)$

مثال ← جدول توزیع توام X و Y را بدست آورید؟

X = تعداد شیرها در ۳ بار پرتاب $\qquad S_X = [0, 1, 2, 3]$

Y = تعداد شیرها در پرتاب آخر $\qquad S_Y = [0, 1]$

$S = \{TTT, TTH, THT, HTT, HHT, HTH, THH, HHH\}$

$P(X=0, Y=0) = \frac{1}{8}$

$P(X=1, Y=0) = \frac{2}{8}$

$P(X=2, Y=0) = \frac{1}{8}$

$P(X=3, Y=0) = 0$

$P(X=0, Y=1) = 0$

$P(X=1, Y=1) = \frac{1}{8}$

$P(X=2, Y=1) = \frac{2}{8}$

$P(X=3, Y=1) = \frac{1}{8}$

مستقل نیستند

y \ x	0	1	2	3	
0	1/8	2/8	1/8	0	4/8
1	0	1/8	2/8	1/8	4/8
					1

سوال ← احتمال اینکه $(x \geq 1$ و $y=0)$ را بدست آورید

(۳ حالت دارد)

$P(X=1, Y=0)+P(X=2, Y=0)+P(X=3, Y=0)=$

از روی جدول حل میکنیم

$\frac{2}{8}+\frac{1}{8}+0=\frac{3}{8}$

‹توزیع حاشیه‌ای›

آیا می‌توان از روی جدول توام توزیع X و Y را بدست آورد؟ بله

فرمول

$P(X=x)=\sum_y P(X=x, Y=y)$

$P(Y=y)=\sum_x P(X=x, Y=y)$

۲ حالت دارد

$P(Y=y) \begin{cases} P(Y=0) \to \sum_{x=0}^{3} P(X=x, Y=0) = P(X=0, Y=0)+P(X=1, Y=0)+P(X=2, Y=0)+P(X=3, Y=0)=\frac{1}{8}+\frac{2}{8}+\frac{1}{8}+0=\frac{4}{8} \\ P(Y=1) \to \sum_{x=0}^{3} P(X=x, Y=1) = P(X=0, Y=1)+P(X=1, Y=1)+P(X=2, Y=1)+P(X=3, Y=1)=\frac{4}{8} \end{cases}$

Y	0	1	
P(Y=y)	4/8	4/8	1

x	0	1	2	3	
P(X=x)	1/8	3/8	3/8	1/8	1

Scanned with CamScanner

نکته ← از روی توزیع توام توزیع حاشیه ای بدست می آید ولی عکس آن ممکن نیست مگر آنکه دو متغیر مستقل هم باشند

سوال ← اگر x و y دو متغیر مستقل باشند توزیع حاشیه آن را بنویسید

$P(X=0 \quad Y=0) = P(X=0)\,P(Y=0)$

$\frac{1}{16} = \frac{1}{8} \times \frac{1}{2}$

شرط استقلال

در توزیع توام $P(A \cap B) = P(A) \cdot P(B)$

یا چند متغیره $P(X=x, Y=y) = P(X=x) \cdot P(Y=y)$

☆ مستقل هستند

چون احتمال ها داخل جدول را با حاصل ضرب احتمالات متناظر بیرون هستند - و کسانی که احتمالات بزرگتر از 1 هستند و جمع احتمالات 1 هست

x/y	۰	۱	۲	۳	
۰	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{3}{16}$	$\frac{1}{16}$	$\frac{1}{2}$
۱	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{3}{16}$	$\frac{1}{16}$	$\frac{1}{2}$

سوال ← احتمال شرطی $P(X=1 \mid Y=1)$ را بنویسید با توجه به جدول غیر مستقل

فرمول $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$

$\frac{P(X=1 \,,\, Y=1)}{P(Y=1 \,,\, X=1)} = \frac{\frac{1}{8}}{\frac{4}{8}} = \frac{1}{4}$

سوال ← احتمال شرطی $P(X=x, Y=y)$ را بدست آورید

$= \frac{P(X=x \,,\, Y=y)}{P(Y=y)}$

سوال ← احتمال شرطی $P(Y=1 \mid X=2)$ با توجه به هر دو جدول هم مستقل هم غیر مستقل بدست آورید.

جدول غیر مستقل $= \frac{P(X=2, \; Y=1)}{P(X=2)} = \frac{\frac{2}{8}}{\frac{3}{8}} = \frac{2}{3}$

جدول مستقل $= \frac{\frac{3}{16}}{\frac{3}{8}} = \frac{1}{2}$

سوال ← احتمال شرطی $P(Y=1 \mid Y=1)$ را بدست آورید با توجه به جدول مستقل

$= \frac{\frac{3}{16}}{\frac{1}{2}} = \frac{3}{8}$

Scanned with CamScanner

سوال ← احتمال

$P(X=x \mid Y=1)$ را بدست آورید

$x \mid y=1$	۰	۱	۲	۳	
$P(X=x$	$\frac{0}{\frac{4}{8}}$	$\frac{\frac{1}{8}}{\frac{4}{8}}$	$\frac{\frac{2}{8}}{\frac{4}{8}}$	$\frac{\frac{1}{8}}{\frac{4}{8}}$	
	۰	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$	① همیشه باید ۱ شود

سوال ← احتمال

$P(X+Y)$ را بدست آورید

$: P(X=0, Y=1) + P(X=1, Y=0) = 0 + \frac{2}{8} = \frac{2}{8}$

سوال ←

$P(X, Y > 2)$ را بدست آورید $\frac{3}{8}$

سوال ←

$X + Y = K$

$P(K=1)$

$S_k = \{1, 2, 3, 4\}$

K	۰	۱	۲	۳	۴	
$P(K=k_j)$	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{3}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	۱

تمرین در خانه ← $P(X-Y=K)$

تمرین در خانه ← $P(X+Y>2)$ یا $P(X>2-Y)$

Scanned with CamScanner

(امید ریاضی) همان میانگین حسابی است

$E(X)$ در جامعه (مقدار ثابت)

$\bar{X} = \frac{1}{n}\sum f_i x_i = \sum f_r(i) x_i$ در نمونه (مقدار متغیر)

$\sum p_i x_i$

میانگین هر عدد ثابت خودش میشود

خواص ها →

$E(a) = a$

$E(aX+b) = aE(X)+b$

$E(X \pm Y) = E(X) \pm E(Y)$

$E(X) = m = E(E(X)) = E(X) = m$

$\bar{X} \approx m$

امید ریاضی ۲ حالت دارد: گسسته = سیگما میگیریم، پیوسته = انتگرال میگیریم

فرمول

$E(X) = \begin{cases} ① \sum_X & \text{(گسسته)} \\ ② \int_X x \cdot f_X(x)\, dx \end{cases}$

مرکزی

$E(X) = \sum_X x \cdot P(X=x)$

$E(X^2) = \sum_X x^2 \cdot P(X=x)$

$E(X^3) = \sum_X x^3 \cdot P(X=x)$

$E(X^4) = \sum_X x^4 \cdot P(X=x)$

گسسته آمار فرمول

$m_1(a) = \frac{1}{n}\sum f_i (x_i - a)^r$

$m_r(\bar{x}) = \frac{1}{n}\sum f_i (x_i - \bar{x})^r$

$m_r(0) = \frac{1}{n}\sum f_i x_i^r$

$\bar{X} \approx m$

$S^2 \approx \sigma^2$

غیرمرکزی

$E(X-m) = 0$

$E(X-m)^2 = V(X) = S^2$

$E(X-m)^3 =$

$E(X-m)^4 =$

⋮

$E(x - E(X)) = E(X) - E(E(X)) = 0$ اثبات کنید

$\sigma^2 = V(X) = E(x - E(x))^2 = E(x^2 - 2X E(x) + E(x)^2)$

$= E(x^2) - 2E(x) \cdot E(X) + E^2(x)$

$= E(x^2) - E^2(x)$

Scanned with CamScanner

سوال ← یک سکه را ۳ بار پرتاب می‌کنیم و $S = ۰\ ۱\ ۲\ ۳$ تعداد شیرها ، به طور متوسط تعداد شیرها در ۳ بار پرتاب چقدر است؟

مقدار مورد انتظار چقدر است؟ ۱٫۵ بار

مقدار مورد انتظار x در ۳ بار پرتاب سکه چقدر است؟

سوال ← یک سکه را ۱ بار پرتاب می‌کنیم انتظار می‌رود چیز با رشیر و چیز با رخط بیاید؟ $x = \frac{1}{2}$

← مثلاً یک سکه را ۵ بار پرتاب می‌کنیم انتظار می‌رود ۲٫۵ بار شیر و ۲٫۵ بار خط رخ دهد.

نحوه‌ی ها امید ریاضی

مثال ← مثال قبلی سکه

x	۰	۱	۲	۳	
$p(X=x)$	۱/۸	۳/۸	۳/۸	۱/۸	۱

بدست آورید؟ ← $E(x) = \sum_{x=0}^{۳} x.p(X=x) = ۰ \times \frac{1}{۸} + ۱ \times \frac{۳}{۸} + ۲ \times \frac{۳}{۸} + \frac{۱}{۸} \times ۳ = \frac{۱۲}{۸} = ۱٫۵$

(گشتاور مرتبه ۲ را حساب کنید) بدست آورید ← $E(x)^۲ = \sum_{x=0}^{۳} x^۲.p(X=x) = ۰^۲ \times \frac{۱}{۸} + ۱^۲ \times \frac{۳}{۸} + ۲^۲ \times \frac{۳}{۸} + ۳^۲ \times \frac{۱}{۸} = \frac{۲۴}{۸} = ۳$

واریانس را حساب کنید $V(X) = (X - ۱٫۵)^۲ = \sum_{X=0}^{۳} (X - ۱٫۵)^۲ . P(X=x) \rightarrow (۰ - ۱٫۵)^۲ . \frac{۱}{۸} + (۱ - ۱٫۵)^۲ \times \frac{۳}{۸} + (۲ - ۱٫۵)^۲ \times \frac{۳}{۸} + (۳ - ۱٫۵)^۲ \times \frac{۱}{۸} = \frac{۶}{۸}$

فرمول دوم ← $E(x^۲) - E^۲(x) = ۳ - (۱٫۵)^۲ = ۰٫۷۵$

Scanned with CamScanner

مثال پیوسته‌ها

امید ریاضی تابع مقابل را بدست آورید؟
واریانس را بدست آورید؟

$$f_X(x) = \begin{cases} 3x^2 & 0 < x < 1 \\ 0 & o.w \end{cases}$$

جواب ←

(امید ریاضی) ← $E(x) = \int_x x \cdot f_X(x)\,dx = \int_0^1 x \cdot 3x^2\,dx = 3\int_0^1 x^3\,dx = 3\left(\frac{x^4}{4}\Big|_0^1\right) = \frac{3}{4}$

فرمول

$$E(x^2) = \int_x x^2 \cdot f_X(x)\,dx = \int_0^1 x^2\, 3x^2\,dx = 3\int_0^1 x^4\,dx = 3\left(\frac{x^5}{5}\Big|_0^1\right) = \frac{3}{5}$$

واریانس

$$V(x) = E(x^2) - E^2(x) = \frac{3}{5} - \left(\frac{3}{4}\right)^2 = \frac{3}{5} - \frac{9}{16} = 0{,}6 - 0{,}56 = 0{,}04$$

واریانس x از طریق مستقیم

$$V(x) = E(x - E(x))^2 = E\left(x - \frac{3}{4}\right)^2 = \int_0^1 \left(x - \frac{3}{4}\right)^2 \cdot 3x^2\,dx$$

〈رابطه میان ۲ متغیر خطی〉 ← کوواریانس (فقط برای گسسته‌ها)

$COV(X,Y)$ را بدست آورید (رابطه خطی بین X و Y)

+ مثبت خطی
- منفی خطی

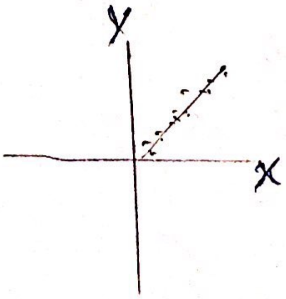
$$CoV(x,Y) = E[(X - E(x))(Y - E(Y))]$$

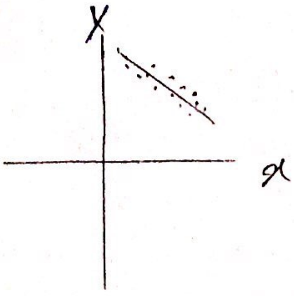
امید ریاضی $E(XY - YE(X) - XE(Y) + E(X)E(Y))$ (ثابت، متغیر، ثابت)

$$= E(XY) - E(Y)E(X) - E(X)E(Y) + E(X)E(Y)$$

$$= E(XY) - E(x) \cdot E(Y)$$

نمودار پراکنش
(یعنی احتمال درکار نیست)





$$Cov(Y,X) = Cov(X,Y)$$

$$= E\left[(X - E(X))\cdot(Y - E(Y))\right]$$

$$= E(X\cdot Y) - E(X)\cdot E(Y)$$

$$Cov(X,X) = V(X)$$

$$Cov(X,a) = 0$$

$$Cov(aX+b,\ cY+d) = a\cdot c\cdot Cov(X,Y)$$

$$V(X \pm Y) = V(X) + V(Y) \pm 2\,Cov(X,Y)$$

$$V(aX \pm bY + C) = a^2 V(X) + b V(Y) \pm 2ab\,Cov(X,Y)$$

اگر مستقل باشند (یعنی ۲ متغیر مستقل یعنی رابطه‌ای بین آنها وجود ندارد)

$$\not\Rightarrow Cov(X,Y) = 0 = E(XY) - E(X)\cdot E(Y)$$

$$E(XY) = E(X)\cdot E(Y)$$

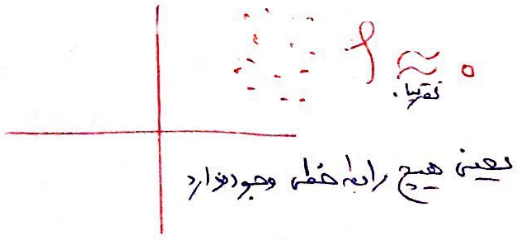
$$\not\Leftarrow X, Y \rightarrow Cov(X,Y) = 0$$

$$\not\Rightarrow P(X = x, Y = y) = P(X=x)\cdot P(Y=y) \Rightarrow E(XY) = E(X)\cdot E(Y)$$

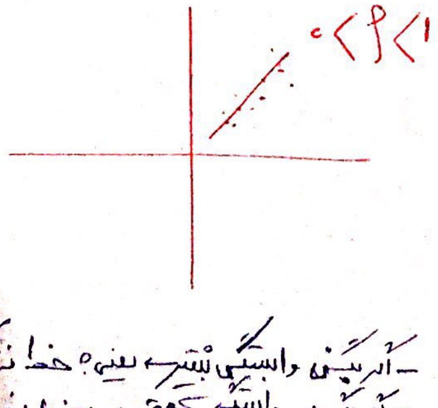
کوواریانس به مقدار وابستگی دارد اگر جامعه خیلی بزرگ یا خیلی کوچک باشد این مقدار هم می‌توانند کوچک یا بزرگ شوند

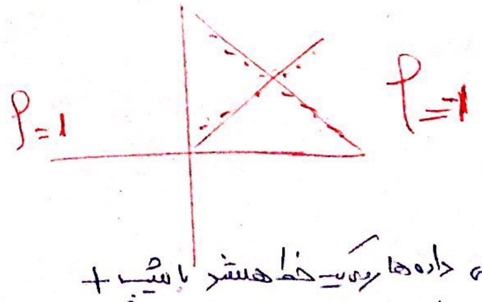
⟨ ضریب همبستگی خطی R ⟩

$$\rho_R = \frac{Cov(X,Y)}{\sqrt{V(X)\cdot V(Y)}} \qquad -1 \le \rho \le 1$$



یعنی هیچ رابطه خطی وجود ندارد





یعنی داده‌ها روی یک خط هستند با شیب +

یعنی داده‌ها روی یک خط هستند با شیب −

- اگر بین وابستگی بیشتر یعنی خط نزدیک‌تر است
- اگر بین وابستگی کم‌تر یعنی از خط دورتر است

Scanned with CamScanner

هرچه R به یک نزدیک‌تر باشد یا ۱- ← داده‌ها به خط نزدیک‌ترند

هرچه R رو به صفر نزدیک‌تر باشد ← یعنی داده‌ها از خط دورترند

ولی COV مشخص نمی‌کند و نزدیکی و دوری را نشان نمی‌دهد

R۰ رو به مقیاس بستگی ندارد

مثلاً ⇐ اگر R یک جامعه شد ۷۵%- یعنی رابطه خطی دارند و خیلی ضعیف است و معکوس هم است

اگر R یک جامعه شد ۰/۳% یعنی رابطه دارند و خیلی ضعیف است

" " ۹۸% خطی است و رابطه آن معکوس است و خیلی به خط نزدیک است

" " ۹۵% + مثبت است و رابطه خطی و به خط نزدیک است

ولی COV این رابطه‌ها را امکان ندارد نشان داد.

۳۹۰

$\frac{4}{8} \cdot \frac{1}{2}$

y \ x	0	1	2	3	
0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	k=0	$\frac{4}{8}$
1	0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	$\frac{4}{8}$
	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	1

سوال احتمالی

۱) k را بدست آورید
۲) مستقل است یا خیر
۳) توزیع حاشیه‌ای x و y را بدست آورید
۴) ضریب همبستگی خطی x و y را محاسبه کنید
۵) $E[(x+1)(y-1)] = ?$
۶) $Cov(3x+1, 1-2y) = ?$
۷) $\rho_{3x+1,\ 1-2y} = ?$
۸) $P(x-y>1) = ?$
۹) $V(x+y) = ?$
۱۰) $P(x \mid y=0) = ?$

$E[(x-1)(y-1)] = ?$

جواب قسمت اول

$\frac{1}{8}+\frac{2}{8}+\frac{1}{8}+k+0+\frac{1}{8}+\frac{2}{8}+\frac{1}{8}=1$

$\frac{8}{8}+k=1 \longrightarrow k=0$

دوم ←

شرط استقلال $P(X=x, Y=y) = P(x).P(y)$

$\frac{1}{8} \neq \frac{4}{8} \times \frac{1}{8}$ ($=\frac{1}{16}$) مستقل نیست

x	0	1	2	3	
P(X=x)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	1

y	0	1	
P(Y=y)	$\frac{4}{8}$	$\frac{4}{8}$	1

$\rho = \frac{Cov(X,Y)}{\sqrt{V(X).V(Y)}} \longrightarrow \frac{\frac{1}{4}}{\sqrt{\frac{3}{4}\cdot\frac{1}{4}}} = \frac{\frac{1}{4}}{\sqrt{\frac{3}{16}}} = \frac{\sqrt{3}}{3}$

$Cov(X,Y) = E(XY) - E(X).E(Y) \longrightarrow 1 - 1.5 \times 0.5 = 0.25 = \frac{1}{4}$

$E(X) = 0\times\frac{1}{8} + 1\times\frac{3}{8} + 2\times\frac{3}{8} + 3\times\frac{1}{8} = 1.5 = \frac{12}{8}$

$E(Y) = 0\times\frac{1}{2} + 1\times\frac{1}{2} = \frac{1}{2}$

$E(X^2) = 0^2\times\frac{1}{8} + 1^2\times\frac{3}{8} + 2^2\times\frac{3}{8} + 3^2\times\frac{1}{8} = 3$

$E(Y^2) = 0^2\times\frac{1}{2} + 1^2\times\frac{1}{2} = \frac{1}{2}$ ⟶

$V(X) = E(X^2) - E^2(X) = 3 - (1.5)^2 = 0.75 = \frac{3}{4}$

$V(Y) = E(Y^2) - E^2(Y) = \frac{1}{2} - (\frac{1}{2})^2 = \frac{1}{4}$

Scanned with CamScanner

$$E(XY) = \sum_{x=0}^{3} \sum_{y=0}^{1} xy \cdot P(X=x, Y=y)$$

$$= \sum_{x=0}^{3} 0 \times x \cdot P(X=x, Y=0) + \sum_{x=0}^{3} 1 \times x \cdot P(X=x, Y=1) =$$

$$= 1 \times 0 \cdot P(X=0, Y=1) + 1 \times 1 \cdot P(X=1, Y=1) + 1 \times 2 \cdot P(X=2, Y=1) + 1 \times 3 \cdot P(X=3, Y=1) =$$

$$= 1 \times 1 \times \frac{1}{8} + 2 \times 1 \times \frac{2}{8} + 1 \times 3 \times \frac{1}{8} = \frac{8}{8} = 1$$

$P(x \mid Y=0) = 0$

اول باید این را بنویسیم

$$\frac{P(X=x, Y=0)}{P(Y=0)} = \frac{P(X=x, Y=0)}{1/2}$$

x	0	1	2	3
$P(x \mid Y=0)$	$\frac{1/8}{1/2}$	$\frac{2/8}{1/2}$	$\frac{1/8}{1/2}$	$\frac{0}{1/2}$

	0	1	2	3	
$P(x \mid Y=0)$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$	0	1

$$Cov(3X+1, 1-2Y) =$$

$$3 \times (-2)\, Cov(X,Y) = -6\left(\frac{1}{4}\right) = -\frac{6}{4}$$

$$Cov(3X+1, 1-2Y) \rightarrow \rho_{3X+1,\, 1-2Y}$$

$$= \frac{Cov(3X+1, 1-2Y)}{\sqrt{V(3X+1) \cdot V(1-2Y)}} = \frac{-\frac{6}{4}}{\sqrt{9 \times \frac{3}{4} \cdot 4 \times \frac{1}{4}}} = \frac{\sqrt{3}}{3}$$

Scanned with CamScanner

توزیع برنولی

$S_X = \{0, 1\}$

$p =$ احتمال موفقیت

$X =$ تعداد موفقیت‌ها در یک بار انجام آزمایش (زمانی که ۲ حالت دارد) $S_X(0,1)$

X	0	1	
$P(X=x)$	$1-p$ (q)	p	1

$p, q > 0$

$p + q = 1$

امید ریاضی $E(X) = 0 \times q + 1 \times p = p$

$E(X)^2 = 0^2 \times q + 1^2 \times p = p$

واریانس → $V(X) = E(X)^2 - E^2(X) = p - p^2 = p(1-p) = pq$

۴ ویژگی آن ۲ حالت است

توزیع دو جمله‌ای ←

$X \sim B(n, p)$

$X =$ تعداد موفقیت‌ها در n بار تکرار مستقل آزمایشی که ۲ حالت دارد

→ تکرار مستقل یعنی: یعنی وقوع یک حالت (حالت) بر حالت‌های دیگر تأثیر ندارد، هر دفعه با دفعه بعد کاملاً مجزا از هم انجام می‌گیرد مانند (نمونه‌گیری با جایگذاری)

یعنی هیچ‌کدام موفقیت ندارد / اصلاً موفقیت ندارد

$S_X\{0, 1, 2, \ldots, n\}$

$P(X=x) = \binom{n}{x} p^x q^{n-x}$ جدول احتمالش با این تابع داریم

مثال ← $(0, 0, 1)$ موفقیت

شکست $(1, 0, 0, 1) \rightarrow = p^2 q^2$ ۴ حالت دارد

باید $\sum_{x=0}^{n} p(X=x) = 1$

اثبات (دو جمله‌ای) $\sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x} = (p+q)^n = 1$

$X = (X_1 + X_2 + \ldots X_n)$

مستقل $X_i \sim B(1, p)$

محاسبه کردن امید ریاضی

$E(X) = E(X_1 + X_2 \ldots X_n)$

$= E(X_1) + E(X_2) \ldots E(X_n)$

حاصل‌جمع $= np$

واریانس $V(X) = (X_1 + X_2 \ldots X_n)$

$= V(X_1) + V(X_2) \ldots V(X_n)$ ← واریانس

$= npq$

(واریانس در جمع پخش می‌شود

واریانس زمانی در جمع پخش می‌شود که کوواریانس‌ها صفر است

یعنی متغیرها از هم مستقل هستند وقتی مستقل باشند کوواریانس‌ها صفر می‌شود)

Scanned with CamScanner

مثال ← یک آزمونی داریم که ۵ تا تست دارد یعنی ۵ گزینه دارد و دارای ۲۰ تا سوال است و سوال‌ها از هم مستقل اند
و کسی که در امتحان شرکت می‌کند به تمام سوالات پاسخ می‌دهد

الف - انتظار داریم کسی که در امتحان شرکت می‌کند به طور متوسط به چند تا سوال جواب صحیح می‌دهد؟

نکته
تفاوت توزیع دوجمله‌ای و برنولی ⇐

دوجمله‌ای است $X \sim p\left(n=20,\ p\ \frac{1}{5}=0.2\right)$

$X \sim B(n,p)$ دوجمله‌ای می‌گوییم
که تکرار تعداد
تعداد پیروزی

$X \sim B(1,p)$
که یکبار تکرار می‌شود

یعنی
(امید ریاضی) → انتظار داریم

یک شخص بی سواد در آزمون که شرکت می‌کند به طور متوسط چند سوال جواب صحیح می‌دهد

$E(X) = n \cdot p = 20 \times \frac{1}{5} = 4$

احتمال اینکه نمره ۴ بیاورد چقدر است؟
نکته (جمع توان باید بشود ۲۰)

$P(X=4)\binom{20}{4} \times (0.2)^{4}(0.8)^{16}$

احتمال درست بودن
احتمال نادرست بودن

باید از جدول ۲۷٪ که در کتاب‌ها داده شده بنویسیم ۳۲٪ حاصل آن → $F_p(4) - F_B(3)$

$= F_B(4) - F_B(3) = 0.630 - 0.411 \approx \%20$

احتمال اینکه این آدم پاس کند چقدر است (۲۰ تا سوال است و هر سوال ۱ نمره است) احتمال چقدر است

$P(X \geq 10) = = 1 - P(x < 10)$

$= 1 - F_B(9) = 1 - 0.997 = 0.003$ جدول کتاب

یعنی از ۱۰۰۰ بار امتحان بدهد انتظار داریم ۳ بار نمره‌اش از ۱۰ بیشتر شود

Scanned with CamScanner

استاد می‌گوید اگر از ۷ نمره بیشتر شوید پاست می‌کنم؟

چقدر احتمال پاس شدن افزایش می‌شود؟

$P(7 \leq x \leq 10)$

$P(7 \leq x < 10)$

$= F_B(\overline{10}) - F_B(\overline{7}) = F_B(9) - F_B(6) =$

$0/997 - 0/913$

از روی جدول کتاب

احتمال اینکه این شخص نمره‌اش از ۱۵ بیشتر شود چقدر است

$P(X > 15)$

$= 1 - P(X \leq 15) = 1 - 1 = 0$

نمره منفی را چقدر در نظر بگیریم که بعد از امتحان میانگین نمره‌اش ثابت صفر (امید ریاضی‌اش صفر شه)

$E(X_i) = 0$ $\quad X_i \sim B(1, 0/2)$

(احتمال انجام شده و شده ۰٫۲)

امید ریاضی

X	$-a$	1	
$P(X = x)$	$0/8$	$0/2$	1

$E(X) = 0$

$= -a \times 0/8 + 0/2 = 0$

$a = \frac{0/2}{0/8} = \frac{1}{4}$ (یعنی عدد صحیح ۴ تا غلط جزمش می‌زند)

شرط بخش می‌شود $\frac{1}{4}$

میزان پراکندگی چقدر خواهد بود؟ واریانس

واریانس $V(X) = npq$

$= 20 \times 0/2 \times 0/8$

احتمال اینکه حداکثر نمره‌اش ۲ باشد ←

$P(X \leq 2) = F_5(2) = 0/206$

Scanned with CamScanner

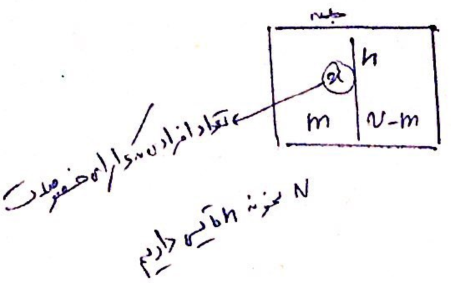
... متغیر ... ۱ ... ۱ ... مستقل و غیر مستقل بودن آنها است

(توزیع فوق هندسی) (در مقایسه با توزیع ۲ جمله ای (مستقل و غیر مستقل بودن است)

فرمول $X \sim HG(N, m, n)$

X : تعداد افرادی که در یک گروه n تایی خصوصیات را دارند

(نمونه n تایی بدون جایگذاری و نمونه از هم مستقل نیستند و غیر مستقل اند)



$$P(X=x) = \frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}} \qquad \sum_x P(X=x) = 1$$

همیشه ۱ می شود

(یک جامعه n تایی داریم n تا دارای خصوصیات هستند m تا این خصوصیات را ندارند
می خواهیم x را بدست؟
نمونه اگر با جایگذاری باشد (دو جمله ای)
اگر بدون جایگذاری) (۱

✓ مثال ← دکتر مهندس (در یک جلسه مثل)

مهندس E_2 پزشک P_3

در یک گروه چهارتایی $n = 4$

$X =$ تعداد مهندس ها

در یک گروه ۴ تایی تعداد مهندس ها چقدر است؟

ضریب کاهش واریانس بر اثر وابستگی داده ها است

واریانس $V(X) = n \cdot \frac{m}{N} \cdot \frac{N-m}{N} \cdot \frac{N-n}{N-1} = 4 \times \frac{2}{5} \times \frac{3}{5} \times \frac{5-4}{5-1}$

امید ریاضی $E(X) = n \cdot \frac{m}{N} = 4 \cdot \frac{2}{5}$

تعداد مهندس ها — جمع دکتر مهندس

جدول توزیع احتمال

X	۱	۲	
$P(X=x)$ احتمال	$\frac{\binom{2}{1}\binom{3}{3}}{\binom{5}{4}}$	$\frac{\binom{2}{2}\binom{3}{2}}{\binom{5}{4}}$	۱
	$\frac{2}{5}$	$\frac{3}{5}$	جمع

$$E(X) = 1 \times \frac{2}{5} + 2 \times \frac{3}{5}$$

$$X \sim HG(n=4,\ m=2,\ N=5)$$

واریانس $4 \times \frac{2}{5} \times \frac{3}{5} \times \frac{5-4}{5-1}$

Scanned with CamScanner

مهم — نکته ← فرق توزیع فوق هندسی و دوجمله‌ای فقط در مستقل و غیر مستقل بودن آنها است

ولی اگر تعداد نمونه زیاد باشد تفاوتی در جامعه‌ها ایجاد نمی‌کند (شرایط ۱)

(و شرایطی است که می‌توانیم توزیع فوق هندسی را میل بدهیم به توزیع دوجمله‌ای) ←

شرایط حدی

$X \sim HG(n, m, N)$

$n \rightarrow 0$ کوچک

$N \rightarrow \infty$ تعداد اعضای جامعه بزرگ

$\frac{m}{N} \rightarrow p$

$\Rightarrow X \sim B(n, p)$

$X \sim B(n, p)$

۱) هرچه داده‌ها هم وابسته‌تر (کوچک‌تر)

مثال ← از بین ۱۰۰۰ متقاضی ۲۰۰ نفر واجد شرایط هستند اگر ۴ نفر برحسب تصادف انتخاب شوند احتمال اینکه ۳ نفر دارای واجد الشرایط باشند چقدر است؟

$N = 1000$
$m = 200$
$n = 4$
$X = 3$

$\rightarrow$

$X \sim HG(4, 200, 1000)$

$P(X = 3) = ? \quad \frac{\binom{200}{3}\binom{800}{1}}{\binom{1000}{4}}$

شرایط حدی می‌گیریم

$\rightarrow \frac{m}{N} = \frac{200}{1000} = 0/2$

$X \sim B(n = 4, p = 0/2)$

$P(X = 3) = \binom{3}{3}(0/2)^3(0/8)^3$

Scanned with CamScanner

پواسون ← حد گرفتن از
یعنی توزیع دوجمله‌ای (۲)

$X \sim B(n,p)$ λ (پواسون)

$p \to 0$

مزدوج: $n \to \infty$ ⟹ $X \sim P(\lambda)$

$np = \lambda$

لاندا

تعداد پیروزی‌هاست
(λ متوسط داده‌های پیروزی)

تعداد اتفاق‌ها X
در یک بازه زمانی یا مکانی
مشخص

$$P(X=x) = \frac{e^{-\lambda}\,\lambda^{x}}{x!}$$

$x \geq 0$

$$\sum_{x=0}^{\infty} P(X=x) = 1$$

امید ریاضی $E(X) = \lambda$

واریانس $V(X) = \lambda$

λ نمی‌تواند هیچ وقت منفی باشد. (چون تعداد متوسط داده‌هاست)

مثال ← در یک کتاب به طور متوسط در صفحه ۳ غلط وجود دارد (۳ غلط در یک صفحه)
احتمال اینکه ۲ غلط وجود داشته باشد چقدر است

$\lambda = \frac{3 \text{ غلط}}{1 \text{ صفحه}}$

$$P(X=2) \to \frac{e^{-3}\,3^{2}}{2!} = 0/423 - 0/199$$

از جدول — از روی جدول (۲-۱)

مثال ← احتمال اینکه ۳ غلط در ۲ صفحه داشته باشد چقدر است؟

$\frac{3 \text{ غلط}}{1 \text{ صفحه}}$ ، λ = ⑥ ، ۲ صفحه

$$P(X=3) \to \frac{e^{-6}\,6^{3}}{3!}$$

$X \sim P(\lambda = 6)$

مثال ← فرض کنید از ۱۰۰۰ نفر یک نفر سرطان داشته باشد. در یک روستا با ۵۰۰۰ نفر جمعیت چقدر احتمال دارد هیچ فردی مبتلا به سرطان نباشد

نفر ۱۰۰۰ ، ۵۰۰۰ — $p = \frac{1}{1000}$ ، $n = 5000$

$N \sim B(n=5000,\ p=\frac{1}{1000})$

$np = \frac{1}{1000} \times 5000 = 0/5 = \lambda$

$P(X=0)$

$$P(X=0) = \frac{e^{-0/5}\,0/5^{0}}{0!} = e^{-0/5}$$

Scanned with CamScanner

< توزیع ۲ جمله ای منفی و هندسی >

X = r امین پیروزی در x امین موفقیت

P = (آزمایشات مستقل)

احتمال پیروزی

مثال ← من یه مقاله می‌فرستم به یک روزنامه با احتمال P مقالات چاپ می‌شود

احتمال اینکه پنجمین مقاله‌ای که این روزنامه می‌فرستم سومین مقاله‌ای است که چاپ می‌شود

$$P(X=x) = ? \rightarrow p\binom{x-1}{r-1} p^{r-1} q^{x-1-(r-1)}$$

$$= \binom{x-1}{r-1} p^{r} q^{x-r}$$

$E(X) = \frac{r}{p}$ امید ریاضی

$V(X) = \frac{rq}{p^2}$ واریانس

مثال عین بالایی ← یه دزد با احتمال P دستگیر می‌شود

احتمال اینکه پنجمین باری که می‌ره دزدی سومین باری است که دستگیر می‌شود

مثال ← احتمال اینکه دزدی در حین دزدی دستگیر شود $\frac{6}{10}$

احتمال اینکه در ۸ هشتمین بار دزدی خود برای چهارمین بار دستگیر شود چقدر است؟

$$X \sim NB(r, p)$$

(احتمال پیروزی)

$r = 4$

$X = 8$

$$P(X=8) = 0{,}6\binom{8-1}{4-1} 0{,}6^{4-1}\; 0{,}4^{8-1-(4-1)}$$

مثال ← اگر احتمال فرزند دختر و پسر با هم مساوی باشد احتمالات زیر را حساب کنید

چهارمین فرزند خانواده‌ای اولین ~~فرزند~~ پسر آن باشد.

دهمین فرزند خانواده‌ای ۴ یا ۵ فرزند پسر باشد

$$P(X=x) = \frac{1}{2}\binom{4-1}{1-1} \frac{1}{2}^{1-1} \times \frac{1}{2}^{4-1-(1-1)}$$

$$= \frac{1}{2} \times 1 \times 1 \times \frac{1}{2}^{3} = \frac{1}{2} \times \left(\frac{1}{2}\right)^{3}$$

Scanned with CamScanner

(توزیع یکنواخت)

مثال یک تیرانداز با احتمال p هدف را می‌زند احتمال اینکه در پنجمین بار تیرش به هدف بخورد چقدر است؟

$p \cdot q^4$

(توزیع یکنواخت)

$X \sim DU(N)$

$P(X = x) = \frac{1}{N}$

مثل پرتاب بدون تاس

$P(X=x)$	x_1	x_2	x_n	
	$1/N$	$1/N$	$1/N$	1

امید ریاضی

$E(X) = \sum_{i=1}^{N} \frac{1}{N} x_i = \frac{1}{N} \sum X_i$

$E(X)^2 = \sum_{i=1}^{N} \frac{1}{N} x_i^2 = \frac{1}{N} \sum X_i^2$

واریانس

$V(X) = E(X)^2 - E^2(X)$

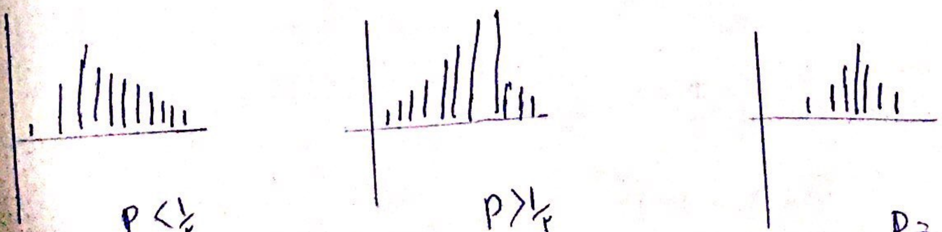
مثال ← امید ریاضی را با توجه به جدول محاسبه کنید؟

X	۱	۳	۵	۷	۹	
	۱/۵	۱/۵	۱/۵	۱/۵	۱/۵	1

$E(X) = \frac{1}{5}(1+3+7+9+5)$

$(E(X))^2 = \frac{1}{5}(1^2+3^2+7^2+9^2+5^2)$

میزان توزیع ۲ جمله‌ای (برای کشیده‌ها)



$p < \frac{1}{2}$ $p > \frac{1}{2}$ $p = \frac{1}{2}$

Scanned with CamScanner

۰٫۲۳

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \; e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2}$$

$x \in R$

$m \in R$

$\sigma > 0$

خواص

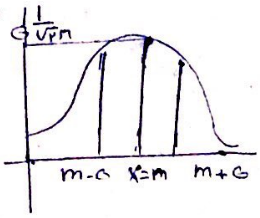
1) $f_X(x) > 0$

2) $\int_{-\infty}^{+\infty} f_X(x)\,dx = 1$

① $X \to \mp\infty \quad y = 0$ مجانب افقی آن محور xها است

② $X = m$ محور تقارن

③ $X = m \quad y = \frac{1}{\sigma\sqrt{2\pi}}$



④ $X = m \pm \sigma$

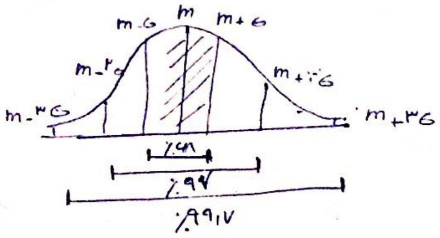
⑤ $E(X) = m$

$V(X) = \sigma^2$

$X \sim N(m, \sigma^2)$

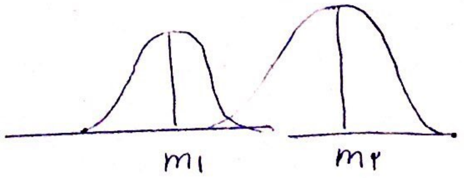
⑥ $F_X(x) = \int_{-\infty}^{x} f_X(x)\,dx$

⑦



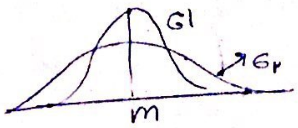
⑧ $m_1 < m_2$

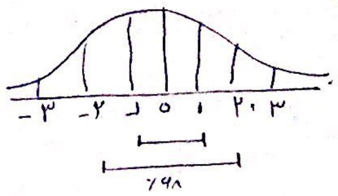
$\sigma = \sigma_1 = \sigma$

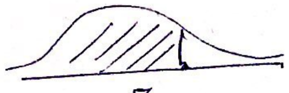


$\sigma_1 < \sigma_2$

$m_1 = m_2 = m$







$Z \sim N(0,1)$

$F_Z(z) = P(Z < z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}\, dx \qquad E(Z)=0 \quad V(Z)=1$

اگر

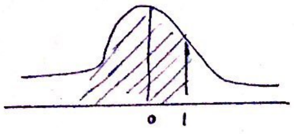
$X \sim N(m, \sigma^2)$

$X = \frac{X-m}{\sigma}$

$\rightarrow X \sim N(0,1)$

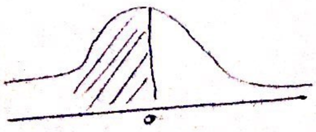
مثال ←

$P(Z<1) = ? \quad 0/8413$



مثال ←

$P(Z \leq 0) = P(Z \geq 0)$



مثال ←

P(Z < -۱) = ?   ۰٫۱۵۸۷

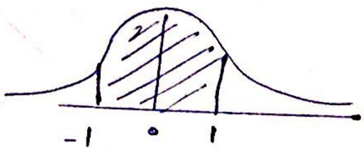
P(Z < -۱) = ۱ - P(Z ≥ -۱)

۱ - P(Z ≤ ۱) = ۱ - ۰٫۸۴۱۳ = ۰٫۱۵۸۷

| P(Z < a) = P(Z > -a) | فقط در توزیع متقارن است

مثال ←

P(-۱ < Z < ۱) = ?



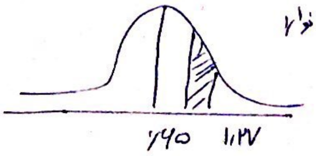
$P(-1 < Z < 1) = F_Z(1) - F_Z(-1)$ تبدیل کرده‌ایم

= ۰٫۸۴۱۳ - ۰٫۱۵۸۷ = ۰٫۶۸۲۶

مثال ←

P(۰٫۶۵ < Z < ۱٫۲۷) ⟹

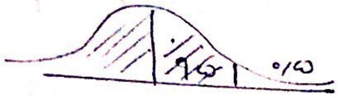
باید تبدیل کنیم $F_Z(1.27) - F_Z(0.65)$



توزیع نرمال استاندارد = ۰٫۸۹۸۰ - ۰٫۷۳۸۹ = ۰٫۱۵۹۱

(استفاده معکوس از جدول توزیع است ندارد)

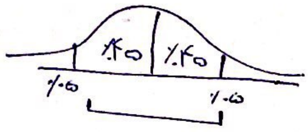
مثال ← ۹۵٪ داده ها زیر چه عددی کمتر است؟



$Z_{0.95} = Z = 1.65$ $Z_{0.95} = Z = 1.65$

$Z_{0.95} \Leftrightarrow P(Z < z) = 0.95$

سوال ← ۹۰٪ داده ها در چه بازه ی متقارن از میانگین قرار دارد؟



$-1.65 = Z_{\%5}$ ٪۹۰ $Z = 1.65$

پس $\Rightarrow$ $Z_{\%95} = Z_{\%05} = 1.65$

خصوصیات ← ۱- تمام توزیع های گسسته و پیوسته هم توانیم تبدیل به توزیع نرمال استاندارد کنیم

۲- حاصل جمع چند توزیع نرمال

مثال ← یک کارخانه داریم که شیشه های آبلیمو را پر می کند متوسط شیشه های آبلیمو ۲۵۰ سی سی است با انحراف معیار ۵ سی سی، احتمال اینکه در این کارخانه شیشه ای را انتخاب کنم که آبلیموی آن از ۲۵۵ سی سی کمتر است چقدر است؟؟

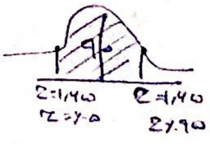
۱- اول باید تبدیل به استاندارد کنیم

$Z = \frac{X-m}{\sigma} = \frac{X-250}{5}$

$P(X<255) = P\left(\frac{X-250}{5} < \frac{255-250}{5}\right) = P(Z<1) = 0.8413$

← ۹۰٪ شیشه آبلیمو در چه بازه ای متقارن از میانگین قرار دارد

$Z = \frac{X-m}{\sigma} \to X = m + Z\sigma \to X = m \pm Z\sigma$

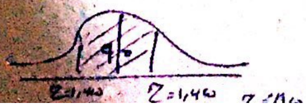


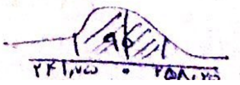
$Z = \frac{X-m}{\sigma}$

$X = m + Z\sigma$

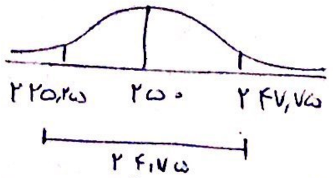
$X = m \pm Z\sigma = 250 \pm 1.65 \times 5$

ادامه ←





← حداکثر تغییر آبگیری (میزان) از کجا تا کجا است؟



۹۹٫۷ داده‌ها انحراف را دارند

$۱۴ \times ۳۵ = ۴٫۹۵$

$۴٫۹۵ \times ۵ = ۲۴٫۷۵$

〈استفاده توزیع نرمال در توزیع دیگر〉

مثال ← اگر یک نمونه ۱۰۰ تایی از این جامعه بگیریم چند تا از شیشه‌ها از ۲۴۵ لیتر بیشتر است؟

$P(X > ۲۴۵) = \longrightarrow P(Z \quad) \quad \frac{۲۴۵ - ۲۵۰}{۵}$

اول باید استاندارد کنیم

$= P(Z > -1)$

$= P(Z < 1)$

$= ۰٫۸۴۱۳$

$۱۰۰ \times ۰٫۸۴۱۳ = ۸۴٫۱۳$

〈تغییر توزیع گسسته با استفاده از توزیع نرمال〉

لاندا

$X \sim P(\lambda)$

امید ریاضی $E(X) = \lambda$

واریانس $V(X) = \lambda$

$Z = \frac{X - m}{\sigma} = \frac{X - \lambda}{\sqrt{\lambda}}$

$\lambda > 10$

$\longrightarrow Z \sim N(0,1)$

$X \sim B(n, P)$

$E(X) = nP$

$V(X) = npq$

$Z = \frac{X - m}{\sigma} = \frac{X - nP}{\sqrt{npq}}$

$\longrightarrow Z \sim N(0,1)$

Scanned with CamScanner

مثال ← فرض کنیم

$X \sim B(n=20,\ p=0.45)$

احتمال آنکه $P(7 \leq x < 14)$ ← ?? را حساب کنید.

$X \sim B(n=20,\ p=0.45)$ احتمال آنکه $P(7 \leq x \leq 14) \longrightarrow$??

$X \sim B(n=20,\ p=0.45)$

$E(X) = np = 20 \times 0.45 = 9$

$V(X) = npq = 20 \times 0.45 \times 0.55 = 4.95 \Rightarrow \sigma = 2.224$ جذر گرفتم

گسسته هست $P(7 \leq x < 14) = F_B(14) - F_B(7) = F_B(14) - F_B(6) = 0.9936 - 0.1299$

$= 0.8637$

برای اینکه گسسته را به پیوسته تبدیل کنم باید تصحیح پیوستگی را انجام دهم

$P(6.5 < x < 14.5)$

$P\left(\frac{6.5-9}{2.224} < Z < \frac{14.5-9}{2.224}\right)$

$= F_Z(2.473) - F_Z(-1.124)$

$= F_Z(2.47) - F_Z(-1.12) = 0.9932 - 0.1314 = 0.8618$

(توزیع نرمال)

$X_i \sim N(m_i, \sigma_i^2) \vdash Y \sim N(\Sigma a_i m_i,\ \Sigma a_i^2 \sigma_i^2)$

$X = \sum_{i=1}^{k} a_i X_i$

X_i مستقل هستند

مثال ← دانشجویی پاس می‌کند که مجموع نمره‌ی پایان ترم و میان‌ترم آن بالاتر از ۱۰ باشد. نمره‌های میان‌ترم کلاس این استاد
از قبل گرفته دارای توزیع نرمال با میانگین ۲ و انحراف معیار ۱
و نمره پایان ترمش با میانگین ۹ و انحراف معیار ۳
احتمال اینکه این دانشجو با استاد پاس کند چقدر است؟؟

$X \sim N(2, 1^2)$

$Y \sim N(9, 3^2)$

$X + Y \sim N(9+2 \; , \; 1^2 + 3^2)$

$N(11, 10) \longrightarrow P(X+Y \geq 10)$

$P\left(Z > \frac{10-11}{\sqrt{10}}\right)$

$P(Z > -0.31) \underset{\text{از جدول}}{\rightarrow} = P(Z < 0.31) = 0.6217$

مثال ← استاد پیشنهاد می‌کند به دانشجویش که نمره میان ترم را ضرب ۵ و نمره پایان ترم را ضرب ۲ می‌کنم
احتمال پاس شدن نمره‌ی بیشتر از ۲۵ آیا حاضری بپذیری یا نه؟

$5 \quad X \sim N(2, 1^2)$

$2 \quad Y \sim N(9, 3^2)$

اول باید احتمال را بدست آوریم ←

$P(5X + 2Y > 25)$

$X + Y \sim N(9+2, \; 1^2 + 3^2)$

$P\left(Z > \frac{25-28}{\sqrt{41}}\right)$

$5X + 2Y \sim N(2 \times 5 + 9 \times 2 \; , \; 5^2 \times 1 + 2^2 \times 3^2)$

$25 + 36$

$P(Z > -0.13)$

$N(28, 41)$

$P(Z < 0.13) \rightarrow 0.5517$

از جدول

< نمونه تصادفی > ←

نمونه تصادفی جمعیت را با متغیر تصادفی X نشان داده که دارای یک توزیع احتمال خاص مانند $f(x)$ که در آن توزیع احتمال جمعیت گویند از این جمعیت یک نمونه n تایی انتخاب میکنیم و اولین عضو x_1 و n امین عضو x_n بنامیم در این صورت مجموع ... که هر کدام از x_i در آن توزیع جمعیت داشته و از یکدیگر مستقل اند را دارای یک نمونه تصادفی میگویند

< پارامتر > ←

هر ویژگی جمعیت را پارامتر نامید که عموماً مقداری نامعلوم و ثابت است

< آماره > ←

هر ویژگی نمونه را آماره گویند و همیشه مقداری معلوم و متغیر است

	پارامتر	آماره
امید ریاضی	$E(X)=m$	$\sim \bar{X}$
	$\hat{m}$	$= \bar{X}$
واریانس	$V(X) \approx \sigma^2$	$\approx S^2U$
	$\hat{\sigma}^2$	$= S^2U$
	p	$\approx \hat{p}$

آماره تابعی از نمونه تصادفی است که به پارامتر مجهول بستگی ندارد و از نمونه ای به نمونه ای دیگر مقدار آن تغییر میکند در حالیکه پارامتر مقدارش همیشه ثابت است

در آمار از یک نمونه تصادفی مانند ... آماره هایی مانند $\bar{x}$ و S^2U محاسبه کرده و از روی آنها در مورد پارامترهای جامعه مانند m و σ^2 استنباط صورت میگیرد

توزیع نمونه ای ←

آماره که توابعی از تصادفی است خود یک متغیر تصادفی است پس میتوان رفتار آن را توسط یک توزیع احتمال توصیف کرد پس هر آماره یک متغیر تصادفی است، توزیع آن را توزیع نمونه ای گویند

تابع محور ←

تابعی است که به نمونه های تصادفی و پارامترهای معلوم و مجهول جمعیت بستگی دارد به طوری که توزیع احتمال آن به پارامتر مجهول بستگی نداشته باشد.

Scanned with CamScanner

مثال — دانشجویی پیش‌بینی می‌کند که مجموع نمره‌هایش در [illegible] از ۱۰ بیشتر [illegible]
[illegible] دارای توزیع نرمال با میانگین ۲ و انحراف معیار ۱
و نمره [illegible] دارای توزیع نرمال با میانگین ۹ و انحراف معیار ۳
احتمال اینکه این دانشجو [illegible] چقدر است؟

$X \sim N(2, 1)$

$Y \sim N(9, 3^2)$

$X+Y \sim N(9+2,\ 1^2+3^2)$

$N(11, 10) \longrightarrow P(X+Y \geq 10)$

$P\left(Z > \frac{10-11}{\sqrt{10}}\right)$

$P(Z > -0.31) = P(Z < 0.31) = \underline{0.6217}$

مثال — استاد پیشنهاد می‌کند دانشجو [illegible] را ضرب در ۵ و نمره [illegible] را ضرب در ۲ کنیم
احتمال اینکه نمره بیشتر از ۲۵ [illegible]؟

$X \sim N(2, 1^2)$

$Y \sim N(9, 3^2)$

اول باید [illegible] ←

$X+Y \sim N(9+2,\ 1^2+3^2)$

$P(5X+2Y > 25)$

$5X+2Y \sim N(2\times5+9\times2,\ 5^2\times1+2^2\times3^2)$

$P\left(Z > \frac{25-28}{\sqrt{61}}\right)$

$N(28, 61)$

$P(Z > -0.384)$

$P(Z < 0.384) \rightarrow 0.6557$

از جدول

۲۸

(قضیه حد مرکزی)

اگر X_i ، n متغیر مستقل با هر توزیع دلخواه و بدون چولگی زیاد داشته باشند که سهم هر کدام از n متغیر ناچیز باشد

آنگاه برای n های بزرگتر $n \geq 10$ آنگاه $Y = \sum_{i=1}^{k} X_i$ دارای توزیع نرمال است به صورت

$Y \sim N(\sum_{i=1} m_i , \sum_{i=1}^{k} \sigma_i^2)$ می‌باشد

$$Y = \sum_{i=1}^{k} X_i$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

$X_i \sim O (m_{(i)}, \sigma^2_{(i)})$ جامعه

$$\sum X_i \sim N(nm, n\sigma^2)$$

$$\sum X_i \sim N(\sum m_i , \sum \sigma_i^2)$$

$$\frac{\sum X_i}{n} = \bar{X} \sim N(m, \frac{\sigma^2}{n})$$

قضیه حد مرکزی $\rightarrow \frac{1}{n} \sum_{i=1}^{k} X_i = \bar{X}$

$X_i \sim O(m_i , \sigma_i^2)$ جامعه

$$\sum X_i \sim N(\sum m_i , \sum \sigma_i^2)$$

نمونه

$X_i \sim O(m, \sigma^2)$ جامعه

$$\sum X_i \sim N(nm, n\sigma^2)$$

$\bar{X} \sim N(m, \frac{\sigma^2}{n}) \rightarrow \frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$ معلوم

توزیع نرمال استاندارد

در مثال زیر ابتدا توزیع X را مشخص کرده و سپس نشان دهید X_i (نمونه تصادفی انتخاب شود)
دارای توزیع جامعه هستند نشان دهید مقدار پارامترهای جامعه مقداری ثابت و مقدار آماره مقداری
متغیر است نشان دهید $\bar{X}$ دارای توزیع نمونه ای نرمال است
و امید ریاضی آن و پارامتر جامعه یکسان است

از ۴ مهره به شماره های ۱، ۲، ۳، ۴ را یک نمونه تصادفی دوتایی انتخاب می کنیم که به طور x_2
با جایگذاری شماره هر مهره انتخاب شده باشد

X	۱	۲	۳	۴	
$P(x)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	۱

X_i	۱	۲	۳	۴	
$P(X_i = x)$	$\frac{4}{16}$	$\frac{4}{16}$	$\frac{4}{16}$	$\frac{4}{16}$	۱

$$E(X) = 1 \times \frac{1}{4} + 2 \times \frac{1}{4} + 3 \times \frac{1}{4} + 4 \times \frac{1}{4} = 2.5$$

امید ریاضی

$$E(X^2) = 1^2 \times \frac{1}{4} + 2^2 \times \frac{1}{4} + 3^2 \times \frac{1}{4} + 4^2 \times \frac{1}{4} = \frac{30}{4}$$

واریانس

$$V(X) = E(X^2) - E^2(X) = \text{———} - \left(\frac{5}{2}\right)^2 = \frac{5}{4}$$

E(

(X_1, X_2)		احتمال انتخاب	$\bar{X}$
۱	۱	$\frac{1}{16}$	۱
۱	۲	$\frac{2}{16}$	۱/۵
۱	۳	$\frac{1}{16}$	۲
۱	۴	$\frac{2}{16}$	۲/۵
۲	۱		۱/۵
۲	۲		۲
۲	۳		۳/۵
۲	۴		۳
۳	۱		۲
۳	۲		۲/۵
۳	۳		۳
۳	۴		۳/۵
۴	۱		۲/۵
۴	۲		۳
۴	۳		۳/۵
۴	۴		۴

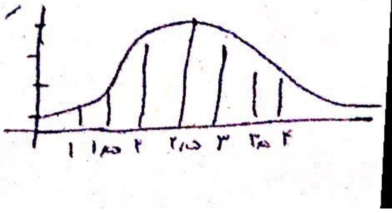
$\bar{X}$	۱	۱/۵	۲	۲/۵	۳	۳/۵	۴
$P(\bar{X} = \bar{x})$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	$\frac{1}{16}$

$$E(\bar{X}) = 2.5 = 1 \times \frac{1}{16} + 1.5 \times \frac{2}{16} = \dots$$

$$E(\bar{X}^2) = 1^2 \times \frac{1}{16} + (1.5)^2 \times \frac{2}{16} \dots$$

$$V(\bar{X}) = \frac{5/4}{2} = \frac{5}{8}$$

خواهیم [illegible] هستش



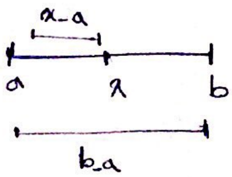
ص ۳۹

پیوسته (فصل ۳) [illegible]

$$X \sim U(a, b)$$

$$a < b$$

X انتخاب یک عدد معین b,a



$$f_X(x) = \begin{cases} \frac{1}{b-a} & a \leq x < b \\ 0 & o.w \end{cases}$$

$$F_X(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x < b \\ 1 & a > b \end{cases}$$

امید ریاضی $E(X) = \frac{a+b}{2} \to ??$

واریانس $V(X) = \frac{(b-a)^2}{12} \to ??$

$$E(X) = \int x . f(x) dx = \int_a^b x \times \frac{1}{b-a} dx = \frac{1}{b-a} \frac{x^2}{2} \Big|_a^b \frac{1}{b-a} \times \frac{b^2 - a^2}{2} = \frac{a+b}{2}$$

$$E(X)^2 = \int x^2 . f(x) dx = \int_a^b x^2 \times \frac{1}{b-a} dx = \frac{1}{b-a} \frac{x^3}{3} \Big|_a^b \frac{1}{b-a} \times \frac{b^3 - a^3}{3} = \frac{a^2 + ab}{3}$$

$$V(x) = E(x)^2 - E^2(x) = \frac{a^2 + ab + b^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}$$

Scanned with CamScanner

‹توزیع نمونه‌ای واریانس› S_U^2 حالت (۱)

قضیه ← اگر $x_1, x_2, \ldots, x_n$ یک نمونه تصادفی است از جمعیت نرمال با میانگین m و σ^2 و X_i ها مستقل از باتوجه به استقلال x_i ها می‌خواهیم قضیه را ثابت کنیم (آنگاه)

الف) $\frac{(n-1)S_U^2}{\sigma^2} \sim \chi^2(n-1)$

ب) ⟸ $\bar{X}$ و S_U^2 از هم مستقل

$S_U^2 \quad X_i \sim N(m, \sigma^2)$

X_i مستقل از

الف → $\frac{(n-1)S_U^2}{\sigma^2} \sim \chi^2(n-1)$

ب → $\bar{X}$ و S_U^2 مستقل از ← یعنی روی هم تاثیر ندارد

فرمول S_U^2 ←

$$S_U^2 = \frac{1}{n-1} \sum (X_i - \bar{X})^2$$

مستقل از ← یعنی توضیحات ۲ تا با هم ربط ندارد و وقوع یکی بر روی احتمال دیگری تاثیر ندارد و احتمال ها را عوض نمی‌کند

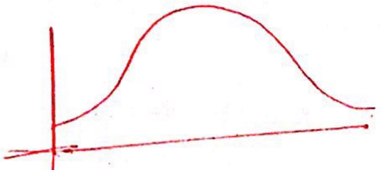
‹توضیح کای دو›

اگر $Z_i \sim N(0,1)$ و از هم مستقل باشند (نرمال استاندارد)

$y \sim \chi^2(n) \leftarrow y = \sum_{i=1}^{n} Z_i^2$

$Z_i \sim N(0,1)$
$y = \sum_{i=1}^{n} Z_i^2$ → $y \sim \chi^2(n)$

شکل کای دو



نکته

۱- توزیع آن نامتقارن است و سمت راست کشیده است

۲- هرچه n زیادتر باشد یعنی درجه آزادی آن بیشتر است و به سمت نرمال می‌رود (توزیع نرمال) غیراستاندارد

توزیع نرمال غیراستاندارد ← چون مرکزش صفر نیست

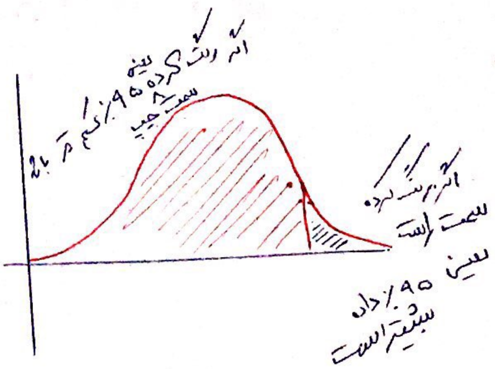
مثال

$y = \chi^2(n) \longleftrightarrow P(Y > \textcircled{y}) = 0.05$

%۹۵

$Y \sim \chi^2(n)$

(یعنی در توزیع کای دو با n درجه آزادی یعنی ۹۵٪ داده‌ها قبل آن است)



$\chi^2(n) - y$

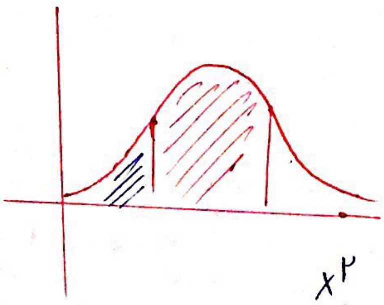
%۹۵

یعنی چی ←

$Y = \chi^2(n) \longrightarrow ??$

$P(Y < y) = 0.95$

$Y \sim \chi^2(n)$



مثال ← از روی جدول صدک ۱۰ توزیع کای دو را پیدا کنید.

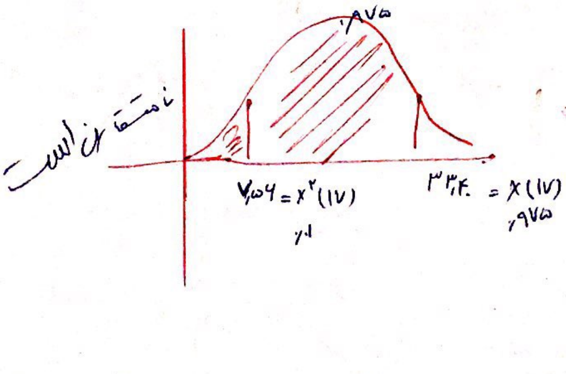
(باید از روی جدول کتاب یا جدولی که فرستاده حساب کنم)

$\chi^2_{0.1}(n) = y = 10.86$

* (استفاده معکوس از جداول خیلی مهم است) *

مثال ←

اگر $Y \sim \chi^2(17)$ آنگاه زیر را حساب کنید ??



$Y \sim \chi^2(17)$

$P(\underset{\chi^2_{0.01}(17)}{7.56} < Y < \underset{\chi^2_{0.975}(17)}{33.40}) \longrightarrow ??$

$= 0.97 - 0.01$

در برخی از کتاب‌ها
اگر سمت راست باشد
۷٫۵۶ – ۳۳٫۴۰
اگر سمت چپ باشد
۳۳٫۴۰ – ۷٫۵۶

Scanned with CamScanner

مثال — یک جمعیت نرمال دارای واریانس ۶ است اگر نمونه‌ای تصادفی ۲۵ تایی از این جمعیت انتخاب، احتمال اینکه واریانس نمونه بیشتر (بین) ۳٫۴۵ تا ۱۰٫۷۵ باشد چقدر است.

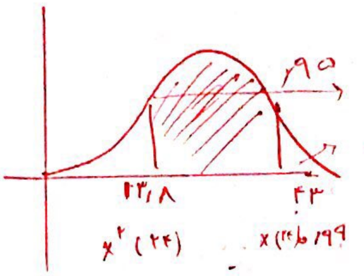
داده‌ها → $X \sim N(m, 6)$ $\quad P(3.45 < S_U^2 < 10.75) =$

$n = 25$

جواب → $= P\left(3.45 \times \frac{(25-1)}{6} < \frac{(n-1)S_U^2}{\sigma^2} < \frac{10.75(25-1)}{6}\right)$

$= P(13.8 < \chi^2(24) < 43) =$

از روی جدول داریم



$= F(43) - F(13.8) = 0.99 - 0.05$

حالت سوم — میانگین و واریانس آن مجهول ←

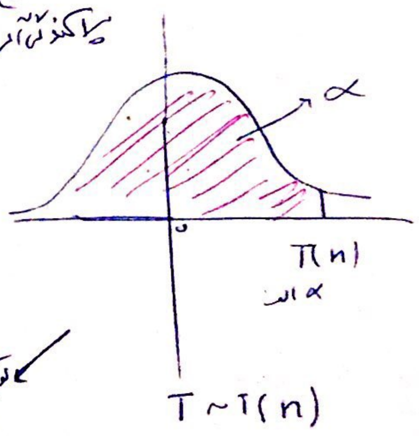
$\bar{X} \to x_i \sim N(m, \sigma^2)$

۳ـ حالت سوم (توزیع T) متقارن است
مرکز آن صفر است
میانگین آن صفر است
مثبت: توزیع نرمال یخ‌تر است
یا کشیدگی آن بیشتر است

σ مجهول

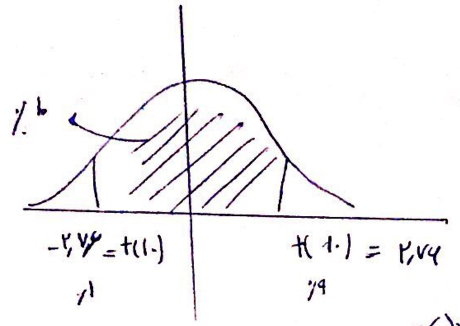
$$\frac{\bar{x} - m}{\frac{s}{\sqrt{n}}} \sim t(n-1)$$

$T(n) \Longleftrightarrow P(T < t) = \alpha$



Scanned with CamScanner

مثال ← صدک دهم در توزیع T با ۱۰ درجه آزادی را بدست آورید



$T \sim T(10)$

بستگی به جدول کتاب دارد

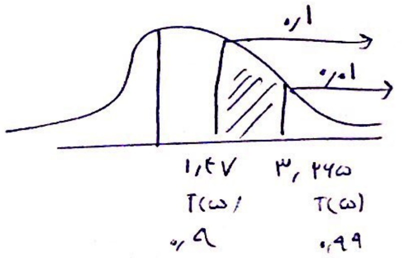
مثال ← اگر T دارای توزیع T و با ۵ درجه آزادی باشد احتمال را بدست آورید

$T \sim t(5)$

$P(1.47 < X < 3.365) \longrightarrow ??$

$P(1.47 < X < 3.365) =$

$F = (-3.365) - F(1.47)$
$\quad\; t(5) \qquad\qquad t(5)$



مثال

مثال ← نمرات یک کلاس از دانشجویان دارای توزیع نرمال با میانگین ۱۵ است. از این کلاس یک نمونه ۲۰ تایی انتخاب کنیم و مشاهده کنیم انحراف استاندارد نمرات ۴٫۲۸ است. احتمال اینکه میانگین نمرات از ۱۷ بیشتر باشد چقدر است؟

داده مسئله

$N(15, \sigma^2)$

$n = 20$

انحراف معیار $S = 4.68$

$P(\bar{X} > 17)$

مجهول است

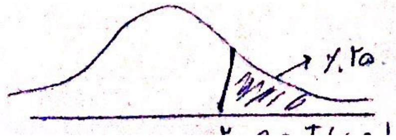
جواب →

$$P\left(\frac{\bar{X} - 15}{\frac{4.68}{\sqrt{20}}} > \frac{17 - 15}{\frac{4.68}{\sqrt{20}}}\right)$$

$$= P(T_{1} > 2.09)$$

$$= 1 - P(T_{(19)} \leq 2.09)$$

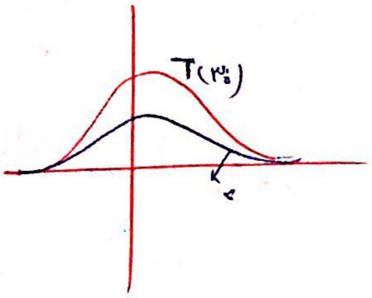
$$= 1 - F_{(19)}(2.09) = 1 - 0.975$$



Scanned with CamScanner

نکته ← اگر n درجه آزادی در توزیع T افزایش پیدا کند توزیع T به سمت نرمال استاندارد میل می‌کند

نکته ← زمانی که از (۳۰) بیشتر باشد =
اگر در جدول T تعداد داده از ۳۰ بیشتر از T استفاده نکنیم
از نرمال استاندارد باید استفاده کنیم



حالت چهارم

نرمال حالت چهارم

$\bar{X}_1 \sim N(m_1, \frac{\sigma_1^2}{n_1})$

$\bar{X}_2 \sim N(m_2, \frac{\sigma_2^2}{n_2})$

مستقل اند σ_1^2, σ_2^2

زمانی که توزیع نمونه‌ای میانگین‌ها وقتی واریانس معلوم است و در بار دوم زمانی که مجهول است و با هم برابر هستند کاری نداریم

$\bar{X}_1 , \bar{X}_2 \sim N(m_1 - m_2 , \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$

$$\frac{\bar{X}_1 - \bar{X}_2 - (m_1 - m_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim (0,1)$$

تا اینجا نمونه‌ها مستقل

برآورد ← استنباط آماری تجزیه و تحلیل و نتیجه‌گیری از بررسی‌ها آماری قسمتی از جمعیت که نتایج آن قابل تعمیم به کل جمعیت باشد استنباط آماری گویند

که دارای ۲ بخش مهم است:
- برآورد پارامتر مجهول
- آزمون فرض در مورد پارامتر مجهول

برآورد پارامتر مجهول به دو صورت انجام می‌گیرد:
- برآورد نقطه‌ای = تقریب یک پارامتر با یک مقدار خاص
- برآورد فاصله‌ای = احتمال قرار گرفتن پارامتر جامعه در یک فاصله خاص

آزمون‌های فرض ← راستی‌آزمایی یک ادعا در مورد پارامترها و یا توزیع یک جمعیت یا یک متغیر تصادفی را آزمون فرض گویند

برآورد نقطه‌ای ← در این روش آماره‌ای به عنوان تقریبی از یک پارامتر مجهول ارائه می‌دهیم

فرض کنیم $T = T(X_1, \ldots, X_n)$ آماره‌ای باشد که از روی نمونه‌ی ما انتخاب شده برای تخمین پارامتر مجهول انتخاب شده است

چنانچه گفته شد T یک متغیر تصادفی است و دارای یک توزیع نمونه‌ای است، T یک برآورد پارامتر گویند و عدد بدست آمده آن برآوردگر آماری گویند

$$T = t(X_1, \ldots, x_n)$$

$$t = T(X_1, \ldots, x_n)$$

Scanned with CamScanner

برآوردهٔ نقطه‌ای باید ۳ خصوصیت داشته باشد

① ناأریبی ← T را یک برآوردگر نااریب گویند اگر امید ریاضی T برابر با پارامتر مجهول باشد $E(T=m)$

② کارایی (هرچه واریانس کوچکتر کارایی آن بیشتر)

③ سازگاری و هم کارایی در احتمال

$E(T)=m$

← نااریبی

$T_1 \quad T_2$

کارایی ← $V(T_1)<V<(T_2)$

$E(T)=m$

برآورد نقطه‌ای ←

هرچه مقدار برآوردگر به پارامتر مجهول نزدیک باشد بهتر است

می‌دانیم مقدار برآوردگر (T) تحت تاثیر نمونه تصادفی انتخاب شده خواهد بود

پس قابل اعتماد نیست به همین دلیل یک فاصله مانند (L,V)

را به عنوان تقریبی از پارامتر مجهول معرفی می‌کنیم که احتمالاً دربرگیرد پارامتر مجهول را در برمی‌گیرد

اگر احتمال قرار گرفتن پارامتر مجهول در این فاصله $1-\alpha$ باشد (L,V) را یک فاصله اطمینان $1-\alpha$ درصد گویند که در آن آلفا بیانگر میزان خطا است

آزمون های فرض آماری ←

آزمون های فرض آماری: فرض آماری یک ادعا در مورد توضیح و یا پارامتر مجهول یک جامعه است که ممکن است درست یا نادرست باشد. مثلا شخصی ادعا کند که میانگین جامعه ای برابر ۷ است می توان این ادعا را با استفاده از آزمون های فرض آزمون کرد.

آزمون های فرض های دارای دو فرض هستند یکی رو فرض H_0 نامیده که این فرض عموماً شامل خلاف ادعا می باشد و فرض H_1 را فرض مقابل نامیده که دقیقاً خود ادعاست.

$H_0 \longrightarrow$ خلاف ادعا (عموماً)

$H_1 \longrightarrow$ ادعا

نکته ← تساوی همیشه در فرض H_0 قرار می گیرد

نکته ← در آزمون فرض ها منظور رد یا قبول فرض هاست با توجه به نمونه گرفته شده رد یک فرض به معنای آن است که از روی نمونه ای که گرفته شده این فرض با قاطعیت رد شده و قبول یک فرض یعنی در نمونه ی بدست آمده دلیلی برای رد فرض نسبت وجود ندارد

< ناحیه ی بحرانی و آماره ی آزمون >

اگر T یک آماره تعریف شده براساس مقادیر مشاهده شده برای رد یا قبول فرض آماری باشد آن را آماره ی آزمون نامیم و ناحیه ای که در صورت قرار گرفتن T در آن فرض H_0 رد می شود را ناحیه ی بحرانی نامیده و این ناحیه را با C بزرگ نمایش می دهند.

< خطا های آزمون >

α رد ناحق H_0 ، B ، B^*

۱_ H : $\alpha\ p($ رد $H_0 \mid H_0$ درست $)$

رد ناحق

۲_ B : $p($ رد $H_1 \mid$ درست $H) = p($ قبول $H_0 \mid$ درست $H_1)$

H_1 رد ناحق

۳_ H_0 : $B^* = p($ رد $H_0 \mid$ درست $H_1) = 1 - p($ رد $H_1 \mid$ درست $H_1) = 1 - B$

Scanned with CamScanner

★ مثال ← (ناحیه بحرانی) امتحان

داروی استاندارد در ۲۵٪ موارد کارا است و شخصی دارای را معرفی نموده است که ادعا دارد که میزان کارایی آن ۵۰٪ است با تعریف ناحیه‌ی بحرانی به صورت $C=\{x|X\geq 9\}$ از یک نمونه‌ی ۲۰ تایی که در آن $X\sim B(n,p)$ میزان خطای اول را حساب کنید

داده شده

$C=\{x|x\geq 9\}$

نمونه $n=20$

$X\sim B(n,p)$

$\alpha=?$

$\beta^{*}=$

جواب →

$\alpha=P(\text{رد } H_0 \mid \text{درست } H_0)=P(X\geq 9 \mid p=\frac{1}{4})$

$X\sim B(n=20,\ p=\frac{1}{4})=P(X\geq 9)=1-P(X<9)$

$=1-F_B(18)=1-0{,}9591=0{,}04=4$ از روی جدول

$\beta=P(\text{رد } H_1 \mid \text{درست } H_1)=P(x<9 \mid p=\frac{1}{2})$

$X\sim B(n=20,\ p=\frac{1}{2})=P(X\leq 8)=$

$=F_B(8)=0{,}2517$ از روی جدول

مثال ← همانند مثال بالا ناحیه بحرانی ۸ را حساب کنید ← $C=\{x|X\geq 8\}$

$\alpha=P(\text{رد } H_0 \mid \text{درست } H_0)=P(X\geq 8 \mid p=\frac{1}{4})=$

$X\sim B(n=20,p\frac{1}{4})$

$=P(X\geq 8)=1-P(X<8)=1-F_B(7)=0{,}1018$

$\beta\to P(\text{رد } H_1 \mid \text{درست } H_1)=P(X<8 \mid p=\frac{1}{2})=P(X\leq 7)=F_B(7)$

$X\sim B(n=20,p\frac{1}{2})$

$=0{,}316$

Scanned with CamScanner

توضیح ناحیه رد دو نوع خطا ها رابطه معکوس تری نشان می دهد همان طور که دیدیم می شود افزایش خطا نوع اول باعث کاهش نوع دوم شد و برعکس ....

بنابر بهترین حالت حالتی خواهد بود که میزان این دو خطا دارای تعادل معقولی باشد

★ نکته ← در آزمون های آماری سعی می شود با قرار دادن یک حداکثر برای α یک مقدار حداقل برای B بدست آوریم

★ مثال ← امتحان

اگر x دارای توزیع نرمال با میانگین m و واریانس $X \sim (m, 4)$ و برای انجام آزمون زیر یک نمونه ۲۵ تایی از جامعه انتخاب کنیم مقدار c کوچک در ناحیه بحرانی را در ناحیه $C = \{x \mid \bar{x} > c\}$ با $\alpha = 0/01$ را بدست آورید و سپس مقدار B را محاسبه کنید با فرض

$$\begin{cases} H_0 = m_0 \\ H_1 = m_1 \end{cases}$$

داده مسئله ↓

$X \sim N(m, 4)$

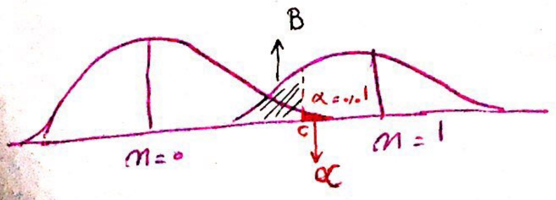
$n = 25$

$c = ?$

$C = \{x \mid \bar{x} > c\}$

$\alpha = 0/01$

$$\begin{cases} H_0 : m = 0 \\ H_1 : m = 1 \end{cases}$$



جواب مسئله →

$X \sim N(0, 4)$

$\bar{X} \sim N(0, \frac{4}{25})$

$$\alpha = P(H_0 \text{ رد} \mid H_0 \text{ درست}) = P(\bar{X} > c \mid m = 0) =$$

$$= P\left(\frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}} > \frac{c - 0}{\frac{2}{5}}\right) =$$

$$P\left(Z > \frac{c}{2/5}\right) = 0/01$$

$$P\left(Z \leq \frac{c}{2/5}\right) = 1 - 0/01$$

$$P\left(Z \leq \frac{c}{2/5}\right) = 0/99$$

$$0/8389 = Z_{.99} = \frac{c}{2/5}$$

$$C = 0/3355$$

مثال ۵ ←

همانند مثال قبل و با داده‌ها مثال قبل $\alpha = 0/1$ را حساب کنید و β را محاسبه کنید

$$\beta = P(\text{رد } H_1 \mid H_1 \text{ درست}) = P(\bar{X} \leq 0/516) = P\left(\frac{\bar{x}-1}{\frac{4}{25}} \leq \frac{0/56-1}{\frac{2}{5}}\right)$$

$$= P(Z < -1/22) = 0/1112$$

$$P\left(Z > \frac{c}{\frac{2}{5}}\right) = 0/1$$

$$P\left(Z \leq \frac{c}{\frac{2}{5}}\right) = 0/1$$

$$P\left(Z \leq \frac{c}{\frac{2}{5}}\right) = 0/9$$

$$1/29 = Z_{0/9} = \frac{c}{\frac{2}{5}}$$

$$C = 0/516$$

$$C = 0/516$$

با کاهش β ← α بزرگ می‌شود و برعکس

(انواع فرض‌های آماری)

فرض‌های آماری به دو دسته تقسیم می‌شوند: ۱. مرکب ۲. ساده

۱ ← مرکب: فرضی که طی آن توزیع جامعه مشخص نمی‌شود (معلوم نیست) مانند ←

$\sigma^2 > 2$
$m \neq 0$
$\sigma^2 < 1$

۲- ساده: فرضی که طی آن توزیع جامعه مشخص می‌شود

مانند ←

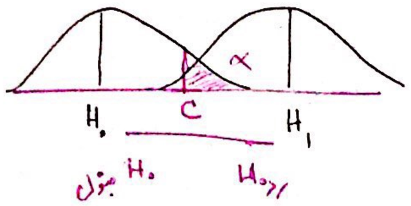
$m = 0$
$m = 1$

Scanned with CamScanner

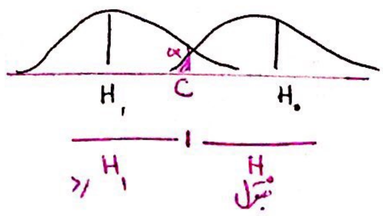
ص ۲۶

<انواع آزمون های آماری>

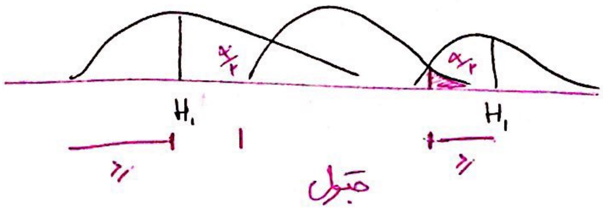
۱- مقادیر بزرگ آماره آزمون باعث رد فرض H_0 می‌شود (یک طرفه راست)



۲- مقادیر کوچک آماره آزمون باعث رد فرض H_0 می‌شود (یک طرفه چپ)



۳- (آزمون ۲ طرفه) مقادیر خیلی بزرگ یا خیلی کوچک باعث رد فرض H_0 می‌شود (دو طرفه)



<آماره آزمون>

همان تابع محوری است به شرط H_0 همان آماره آزمون است

★ مثال امتحان ←

نمونه ی تصادفی از پرونده های شرکتی نشان می‌دهد که سفارشات برای قطعه معینی از ماشین ها به ترتیب ۱۰, ۱۲, ۱۹, ۱۴, ۱۵, ۱۸, ۱۱, ۱۳ بایگانی شده است

اگر تعداد روزهای بایگانی از توزیع نرمال پیروی کند آیا از سطح معنی‌داری ۰/۰۱ میتوان ادعا کرد که میانگین زمان بایگانی چنین سفارشاتی از ۱۰/۵ روز بیشتر است

$X \sim N(m, \sigma^2)$

σ^2 مجهول

$\bar{X} \sim N(m, \frac{\sigma^2}{n})$

$\frac{\bar{X} - m}{\frac{s}{\sqrt{n}}} \sim t(n-1)$

نمونه $n = 8$

$\sum_{i=1}^{8} X_i = 112$

$\sum_{i=1}^{8} X_i^2 = 1614$

$\bar{X} = \frac{112}{8} = 14$

$S_u^2 = \frac{1}{n-1}\left(\sum X_i^2 - \frac{1}{n}(\sum X_i)^2\right) =$

$\alpha = 0.01$

ادعا: $m > 10.5$

$H_0 = m \leq 10.5$

$H_1 = m > 10.5$

$T = \frac{14 - 10.5}{\sqrt{\frac{10.284}{8}}} = 3.087$ آماره آزمون

توزیع مقدار ... 0.99 ... $\alpha = 0.01$

★ مثال ← یک فاصله اطمینان ٪۹۹ برای این آزمون بدست آورید

فرمول → $m: \bar{X} \pm \frac{s}{\sqrt{n}} + (n-1)$ ٪۹۹

$14 \pm \sqrt{\frac{10.286}{8}} \cdot$ ۴(۷) ٪۹۹

★ مثال ←

یک تولیدکننده‌ی قطعات پیش ساخته مدعی است که انحراف معیار مقاومت محصولات او برابر ۱۰ کیلوگرم بر سانتی‌متر مربع است
یک نمونه تصادفی ۱۰ تایی از این محصولات نتایج $\bar{X} = 3.12$ و $s^2 = 195$ را بدست می‌دهد
اگر اندازه‌ی مقاومت این محصول دارای توزیع نرمال باشد آیا این نتایج بدست آمده
با ادعای تولیدکننده سازگار و راست سطح معنی‌داری را ۰٫۰۵ بگیرید

داده مسئله

$\sigma = 10$: ادعا

$n = 10$

$\bar{X} = 3.12$

$s^2 = 195$

$X \sim N(m, \sigma^2)$

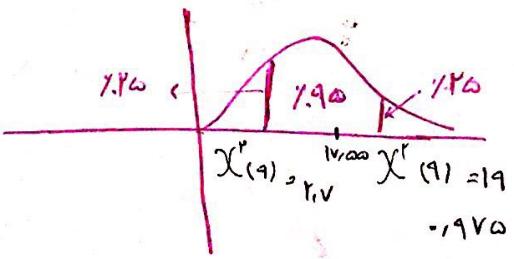
$\alpha \; 0.05$

$H_0 : \sigma^2 = 100$

$H_1 : \sigma^2 \neq 100$

$\longrightarrow \frac{(n-1) \times s^2}{\sigma^2} \sim \chi^2(n-1)$ (کای دو)

$\frac{(10-1) \times 195}{100} = 17.55$



از روی جدول کای‌دو محاسبه

* حال اگر در امتحان گفته شود ٪۹۰ برای واریانس این جامعه را بدست آورید

* و یک فاصله ٪۹۰ برای میانگین این جامعه را بدست آورید

Scanned with CamScanner

۱۴۷

۱- یک فاصله اطمینان برای m میانگین جامعه با واریانس معلوم σ^2 بدست آورید؟

$$P(L < m < U) \quad 1-\alpha$$

دامنه‌ها

$\%(1-\alpha)$

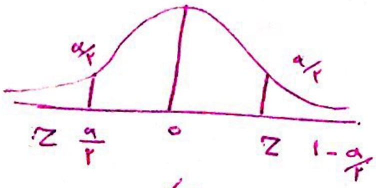
(L, U)

$$P\left(a < \frac{\bar{x} - m}{\frac{\sigma}{\sqrt{n}}} < b\right) \doteq 1-\alpha$$

$$P\left(\underbrace{\bar{x} - b\frac{\sigma}{\sqrt{n}}}_{L} < m < \underbrace{\bar{x} - a\frac{\sigma}{\sqrt{n}}}_{U}\right) = -\alpha$$

$$= \bar{x} - Z\frac{\sigma}{\sqrt{n}} \qquad m < \bar{x} + Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

$$m : \bar{X} \pm \frac{\sigma}{\sqrt{n}} \cdot Z_{1-\frac{\alpha}{2}}$$



چون متقارن است
با هم برابر هست

مثال ←
$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$$

متقارن نیست

تابع محور $P(L < \sigma^2 < U) = 1-\alpha$

واریانس $P(L < \sigma^2 < U) = 1-\alpha$

$$P\left(a < \frac{(n-1)}{\sigma^2} < b\right) = 1-\alpha$$

$$P\left(\frac{(n-1)s^2}{b} \quad \sigma^2 < \frac{(n-1)s^2}{a}\right) = 1-\alpha$$

$$(n-1)s^2 < \sigma^2 < \frac{(n-1)s^2}{\chi^2(n-1)}$$

$1-\alpha/2$ $\qquad$ $\frac{\alpha}{2}$

تفاضل $m_1 - m_2 : \bar{x}_1 \bar{x}_2 \pm \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

$m : \bar{x} + \frac{\sigma}{\sqrt{n}} + (n-1)$

★ مثال ←

نمونه سوال

(تا اینجا برای جلسه دوم)

اگر طول قد کارمندان یک اداره دارای توزیع نرمال است یک فاصله اطمینان ۹۵٪ درصد برای میانگین طول قد کارمندان پیدا کنید در حالتی که نمونه ۵ تایی از بین کارمندان انتخاب و قد آنها برابر ۱۶۰ و ۱۷۰ و ۱۶۵ ، ۱۷۵ ، ۱۸۰ باشد

مجهول σ^2

$X \sim (m, \sigma^2)$

$1-\alpha = 0.95$

$m = ?$

$n = 5$

160, 170, 180, 175, 165

$\sum_{i=1}^{5} x_i = 850$

$\sum_{i=1}^{5} x_i^2 = 144750$

$n = 5$

اول باید اینها را بدست بیاوریم بعد حل میکنیم

$$\bar{X} = \frac{160+170+175+165+180}{5}$$

$$\bar{X} = \frac{\sum_{i=1}^{5} x_i}{5} = \frac{850}{5} = 170$$

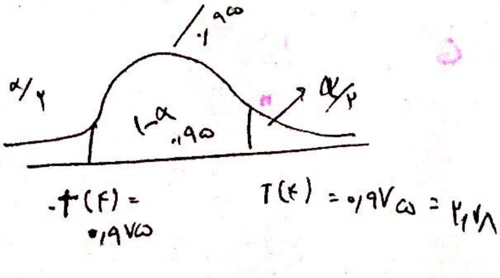
$$s_v^2 = \frac{1}{n-1}\left(\sum x_i^2 - \frac{1}{n}\left(\sum x_i\right)^2\right)$$

$$= \frac{1}{4}\left(144750 - \frac{1}{5}(850)^2\right) = 62.5$$

$$m : \bar{X} \pm \frac{s}{\sqrt{n}}\, t_{1-\alpha/2}(n-1)$$

$$\text{از روی جدول} = 170 \pm \sqrt{\frac{62.5}{5}} \times 2.78 = 160.17$$

$$(160.17 , 179.83)$$



درجه آزادی ۴